

ADA130260

(12)

(UNCLASSIFIED)

AD \_\_\_\_\_

TASK VALIDATION FOR STUDIES ON FRAGMENTED SLEEP  
AND COGNITIVE EFFICIENCY UNDER STRESS

ANNUAL AND FINAL REPORT

Charles Graham, Ph.D.

~~JAN 1983~~  
*November 1982*

Supported by

U.S. Army Medical Research and Development Command  
Fort Detrick, Frederick, Maryland 21701

Contract No. DAMD17-80-C-0075

Midwest Research Institute  
Kansas City, Missouri 64110

DTIC  
S JUL 11 1983  
A

Approved for Public Release; Distribution Unlimited

The findings in this report are not to be construed as an  
official Department of the Army position unless so design-  
ated by other authorized documents.

(UNCLASSIFIED)

DTIC FILE COPY

83 07 08 021

(UNCLASSIFIED)

AD \_\_\_\_\_

TASK VALIDATION FOR STUDIES ON FRAGMENTED SLEEP  
AND COGNITIVE EFFICIENCY UNDER STRESS

ANNUAL AND FINAL REPORT

Charles Graham, Ph.D.

~~1982~~  
*November 1982*

Supported by

U.S. Army Medical Research and Development Command  
Fort Detrick, Frederick, Maryland 21701

Contract No. DAMD17-80-C-0075

Midwest Research Institute  
Kansas City, Missouri 64110

Approved for Public Release; Distribution Unlimited

The findings in this report are not to be construed as an  
official Department of the Army position unless so design-  
ated by other authorized documents.

(UNCLASSIFIED)

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO. AD-A130 260	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) TASK VALIDATION FOR STUDIES ON FRAGMENTED SLEEP AND COGNITIVE EFFICIENCY UNDER STRESS		5. TYPE OF REPORT & PERIOD COVERED Annual Report (4/1/81 - 9/30/82) and Final Report (5/1/80 - 9/30/82)
7. AUTHOR(s) Charles Graham, Ph.D.		6. PERFORMING ORG. REPORT NUMBER MRI No: 2007-E
9. PERFORMING ORGANIZATION NAME AND ADDRESS Midwest Research Institute 425 Volker Boulevard Kansas City, MO 64110		8. CONTRACT OR GRANT NUMBER(s) DAMD17-80-C-0075
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Medical Research and Development Command Fort Detrick, Frederick, Maryland 21701		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62777A, 3E162777A879, AB.005.
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)		12. REPORT DATE November 1982
		13. NUMBER OF PAGES 89
		15. SECURITY CLASS. (of this report)  Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report)  Approved for public release; distribution unlimited		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)  Approved for public release; distribution unlimited		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number)  Human performance, information processing, decision making, Risk-taking behavior, Workload, Fragmented sleep, Stress, Crisis, Error analysis, Computer gaming C3 operations		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number)  A computer gaming approach was used to create a new type of automated per- formance task (STAR). The task unobtrusively measures multiple cognitive skills and risk-taking behavior under various stress and workload conditions. A training manual and protocol were developed, and performance criteria es- tablished. Measurement reliability and performance under different task dif- ficulty levels and crisis conditions were assessed. STAR is shown to be a sensitive task which promises to reliably measure major aspects of human func-		

DD FORM 1 JAN 73 1473 A EDITION OF 1 NOV 55 IS OBSOLETE

tion under a variety of conditions.

UNCLASSIFIED  
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

(UNCLASSIFIED)

**AD**

## TASK VALIDATION FOR STUDIES ON FRAGMENTED SLEEP AND COGNITIVE EFFICIENCY UNDER STRESS

## ANNUAL AND FINAL REPORT

Charles Graham, Ph.D.

July 1983

Supported by

U.S. Army Medical Research and Development Command  
Fort Detrick, Frederick, Maryland 21701

Contract No. DAMD17-80-C-0075

Midwest Research Institute  
Kansas City, Missouri 64110

Approved for Public Release; Distribution Unlimited

The findings in this report are not to be construed as an official Department of the Army position unless so designated by other authorized documents.

(UNCLASSIFIED)

be construed as an  
on unless so design-  
ments.

Accession For

THIS GRA&I

Form TAB

Unclassified

Justification

Distribution/

Availability Codes

Avail and/or

Special

DTIC

COPY

INSPECTED

3

A

## FOREWORD

This document describes the work accomplished by Midwest Research Institute, and the results obtained, in a research project conducted under Contract No. DAMD17-80-C-0075 with the U.S. Army Medical R&D Command during the period May 1, 1980 through September 30, 1982. Dr. Frederick W. Hegge (WRAIR), and Dr. James Stokes (USARIEM) served as COTRs during the first phase of project activity, and Dr. Hegge continued in this capacity for the remainder of the project period.

The Principal and Co-Principal Investigators were Dr. Charles Graham and Mr. Harvey D. Cohen, respectively. Mr. James W. Phelps wrote the computer programs used to implement the performance task described. Ms. Mary M. Gerkovich performed the statistical analyses presented, and Mr. Michael A. Hamilton aided in the conduct of the experiments and the data analyses. Dr. Mary R. Cook originally suggested the concept further developed in this project, aided in the research design, and supervised the statistical analyses. Dr. Sophia S. Fotopoulos provided administrative supervision and support for the project. Each of the above individuals made numerous contributions to the project, and the work presented here represents the combined efforts of all involved. For the protection of human subjects the investigators have adhered to policies of applicable Federal Law 45CFR46.

## TABLE OF CONTENTS

	<u>Page</u>
Summary. . . . .	1
I. Introduction and Objectives. . . . .	3
II. Specific Aims. . . . .	4
III. Phase I: Development of a New Human Performance Task. . . . .	5
A. Initial Considerations . . . . .	5
B. Project Activities . . . . .	7
C. Description of STAR (Strategic and Tactical Assessment Record). . . . .	8
IV. Phase II: Training and Task Evaluation. . . . .	15
A. Development and Testing of a Training Protocol . . . . .	21
B. Establishment of Three Levels of Task Difficulty . . . . .	24
C. Evaluation of Specific Stressors on Performance of STAR. . . . .	50
D. Evaluation of Test-Retest Reliability. . . . .	66
V. Conclusions and Recommendations. . . . .	72
Appendix A - Star Training Protocol. . . . .	76
Appendix B - Some Suggested Areas for Future Research. . . . .	82
Appendix C - Evaluation of Star as an Experimental Tool. . . . .	85

# LIST OF FIGURES

<u>Number</u>	<u>Title</u>	<u>Page</u>
1	Basic Information Display Pattern of the VENTURE Cruiser .	14
2	Volunteer's Informed Consent . . . . .	20
3	Acquisition of Tracking Skills Using the PATH Program. . .	25
4	Change in Mean Time Required to Destroy Xenoids as a Function of Training Trials. . . . .	26
5	Mean Time per Command, A Measure of Overall Performance Efficiency, as a Function of Training Trials . . . . .	27
6	Mean Time Between Long-Range Scans as a Function of Training Trials . . . . .	28
7	Effects of Difficulty Level Factors on Time (Left Axis) and Energy (Right Axis) Utilization. . . . .	34
8	Effects of Star Density and Xenoid Distribution on Manuever and Search Commands . . . . .	39
9	Effects of Star Density and Xenoid Dispersion on Defensive Strategy . . . . .	43
10	Interactions Between Sex and Xenoid Dispersion . . . . .	48
11	Percent Weapon Commands Issued Without Adequate Shield Protection . . . . .	57
12	Mean Shield Level on Entering Enemy Quadrants. . . . .	58
13	Effects of Stressors on Risk-Taking Behavior During Attack Situations . . . . .	59
14	Number of LRS and LRSH Commands With No New Information. .	61
15	Percent Quadrants of New Information per LRS . . . . .	62
16	Percent Noncontiguity of Quadrant Mapping. . . . .	63
17	Mean Time to Extract Information per LRS . . . . .	64

# LIST OF TABLES

<u>Number</u>		<u>Page</u>
1	Command and Control Systems of the VENTURE . . . . .	13
2	Sample Output Print Showing Major Performance Parameters of STAR. . . . .	16
3	Promotion Criteria for VENTURE Captain Level 10 Through 1.	23
4	Dispersion of Enemy Units in 30 Training Galaxies. . . . .	29
5	Characteristics of Test Galaxies Used in Difficulty Level Experiment . . . . .	30
6	Changes in Subjective Ratings as a Function of Difficulty Level. . . . .	31
7	Table of Means Significant F Values, Subjective Ratings. .	32
8	Table of Means Significant Changes in Major Performance Variables. . . . .	33
9	Effects of Task Difficulty on Variables Related to Navigation . . . . .	35
10	Table of Means Significant Variables Related to Navigation . . . . .	36
11	Effects of Task Difficulty on Variables Related to Search Strategy . . . . .	37
12	Table of Means Significant Variables Relevant to Search Strategy . . . . .	38
13	Effects of Difficulty on Variables Related to Offensive Strategy . . . . .	41
14	Table of Means for Significant Variables Related to Offensive Strategy . . . . .	42
15	Variables Showing Significant Main Effects for Sex of Subject. . . . .	45
16	Significant Interactions Between Sex and Xenoid Dispersion . . . . .	46
17	Experimental Design of the Stress Study. . . . .	53
18	Characteristics of Eight Test Missions Used in the Stress Study. . . . .	54

LIST OF TABLES (concluded)

<u>Number</u>		<u>Page</u>
19	Effects of Time and Energy Restrictions on Subjective Ratings. . . . .	55
20	Preliminary Examination of Measures for Nomality and Test-Retest Peliability . . . . .	67
21	Test-Retest Reliability for STAR Measures. . . . .	69

## SUMMARY

There has been a continuing problem in attempting to adequately characterize and evaluate the type of complex cognitive functioning required for efficient C<sup>3</sup> operations. Previous research has been criticized for its use of relatively simple tasks, and simulation exercises have been faulted for their reliance on subjective assessments of performance adequacy. Recent technical advances have only compounded the problem. These advances have decreased the margin for error and at the same time, they have placed additional demands on the human ability to process information, assess risks, and make complex decisions under the pressures of time, ambiguity, and shifting priorities.

A new approach to the assessment of complex human performance is needed; one that focuses on relevant cognitive abilities, and provides objective, operationalized measures of human error and the decision paths that lead to error. The ultimate goal of such an approach should be the development of effective techniques to aid personnel in maintaining cognitive performance efficiency under adverse conditions. The project described in this report has contributed to this ultimate goal through the development of a prototype of the "next generation" of experimental research tasks.

The prototype task developed in this project is called STAR (Strategic and Tactical Assessment Record). STAR is intended to be used as a versatile research tool to evaluate multiple parameters of cognitive efficiency under various environmental, personal, and situational stress conditions. It is a completely automated, complex performance task set in the context of a futuristic war scenario. STAR provides multiple operationalized measures of an individual's visual motor coordination, subjective state, attention, memory, perception, information processing, risk taking behavior, errors and error paths, and both strategic and tactical decision making. One of the unique features of STAR is that all measurement procedures are buried within the context of the operations required to perform the task. STAR also has unique motivational properties built into it, which aid in maintaining high levels of sustained effort and performance involvement over extended periods of time.

The following sections of this document describe the development and testing of STAR during the period May 1, 1980 through September 30, 1982. Project activities have included (a) creation of the task and the computer programs necessary for presentation and data collection; (b) development of an efficient training protocol and preparation of a training manual and task exercise problems; (c) the performance of a study to establish three levels of task difficulty, and to assess cognitive functioning at different levels of task difficulty; (d) the performance of a study to evaluate STAR performance under crisis and noncrisis conditions; and (e) a statistical study to evaluate the test-retest reliability of STAR measurement parameters. The findings reported here show STAR to be a sensitive research instrument which shows promise of reliably measuring major aspects of human function under a variety of relevant conditions. These findings

also demonstrate that future research is necessary to more completely develop this prototype task, and to followup a number of the research directions indicated.

## I. INTRODUCTION AND OBJECTIVES

Three major trends are apparent in the world-wide, military technical environment. First, the introduction of advanced computer and electronics technology is rapidly increasing the complexity of the environment and making it highly machine interactive. Second, recent advances in weapons research and development are changing established concepts of warfare. Vastly greater firepower can now be concentrated in small combat units, and high mobility unit dispersion, rather than massed troop concentration, is becoming the rule. Third, and most important, efficient use of the new technologies at the C<sup>3</sup> level is becoming increasingly dependent on a particular set of human higher-order mental abilities. These include: (1) the ability to assimilate high rates of multi-source, variable priority information; (2) the ability to integrate and use this information on a real time basis; (3) the ability to accurately assess risks and make complex decisions under the pressures of time, ambiguity, and shifting priorities; and (4) the ability to maintain cognitive performance efficiency in the face of sustained performance demands and/or under situational and environmental stress conditions.

These trends lead to the clear need for research designed to evaluate the impact of likely stressors on complex cognitive functioning. Ideally, such research should allow the detailed analysis of human error, and the tracking of error paths in information processing, decision making and risk taking behavior. The ultimate goal of such research should be the development of effective techniques to aid men and women in maintaining cognitive performance efficiency under adverse conditions.

The relevance of previous research to this need has been questioned on a number of counts. The majority of sustained performance studies have examined the effects of extended time spent in a low stimulation environment on the performance of relatively simple and often boring tasks (e.g., radar monitoring). Studies in cognitive psychology have used more demanding tasks; however, such tasks have often been more appropriate to the laboratory assessment of college freshmen than to evaluation of the type of complex, integrated functioning required in C<sup>3</sup> operations. Finally, the utility of full-blown C<sup>3</sup> simulation has been questioned since the measures obtained are often not operationalized, and in many instances, evaluation is dependent on the subjective report of observers.

Such criticisms suggest that a different approach is required. A research protocol needs to be developed that focuses specifically on those cognitive and performance requirements most relevant to a high technology environment. This protocol, or test bed, should provide operationalized measures of relevant performance parameters. It should allow the introduction and manipulation of variable workload requirements as well as situational or environmental stressors. Finally, the protocol should include intrinsically interesting motivational components, and provide both positive and negative consequences based on the actions of the individual.

These considerations led to the decision to create a new computerized, cognitive performance task/scenario that would be patterned after existing computer games, and would incorporate the characteristics described above. The advantage of using this type of basic task activity are: (1) performance of the task could be made to consist almost entirely of information processing, decision making and risk assessment activities; (2) task activities would occur continuously as discrete units of time-limited behavior capable of being quantified and operationalized; (3) high levels of sustained interest have been demonstrated in real world settings with this type of task; and (4) stress factors and reinforcement contingencies can be introduced and altered in a systematic and dynamic fashion.

Thus, the objectives of this research program were to first develop an initial version of the task/scenario, and then to conduct a series of experiments to establish training procedures and performance criteria, and to determine the reliability and sensitivity of the measures derived.

## II. SPECIFIC AIMS

The specific aims of this project were:

- \* Create a complex, cognitive performance task patterned after existing computer games, and incorporating unobtrusive, multiple measures of information processing, decision making and risk taking behavior.
- \* Develop the necessary materials and an efficient protocol for the training of individuals to perform the task.
- \* Test the training procedures on a sample of 6 men and 6 women to establish training performance criteria.
- \* Conduct an experiment on 12 trained subjects to evaluate the test-retest reliability of the task measures.
- \* Conduct an experiment on 12 trained subjects to test the effects of different levels of task difficulty on task performance.
- \* Conduct an experiment on 12 trained subjects to test the effects of time pressure on task performance.
- \* Conduct an experiment on 16 trained subjects to test the sensitivity of the task to sustained performance demands and circadian rhythm effects.
- \* Prepare all materials necessary for an operational human performance testing package.

### III. PHASE I: DEVELOPMENT OF A NEW HUMAN PERFORMANCE TASK

#### A. Initial Considerations

In preparing the proposal for the work described in this report, the staff carefully considered the characteristics of the task to be developed. The first question was whether the task should evaluate the performance of a group or of an individual. It was thought that an individual task would be more generally applicable to a variety of settings, and would allow research and testing to be conducted in the personnel selection and individual difference areas. Once this was accomplished, follow-on activities could focus on the development of a group-interactive task.

1. Properties of the task: It was decided that the task would have four major properties. First, it would include a comprehensive set of operationally-defined measurement parameters. These would provide evaluation of specific sub-areas of human function such as reaction time and visual tracking ability, as well as assessment of more complex integrated functions such as the number and type of resources an individual used in arriving at a particular decision. Additional measures would evaluate multiple parameters of visual motor coordination, subjective state, attention, memory, perception, information processing, risk taking behavior, resource use and allocation, errors and error paths, and both strategic and tactical decision making. Our intent was to create a task that would provide a comprehensive, well-rounded assessment of human performance over a defined unit of time. Time units could then be linked in series to assess performance over any experimental time period of interest.

The second major property was that all the measurement procedures described above would be uniquely buried within the context of the operations required to play a computer game. From the individual's point of view, he or she would be playing a futuristic computer war game. After training, the individual would be left alone in a room with a computer terminal and display scope and simply play the game. No researcher would be present. No one would be overtly measuring reaction time, attention, perception, etc. In other words, the game context would be used to justify and present a wide variety of measures commonly employed in the laboratory setting, but without the interference effects, boredom or performance anxiety often associated with such measurement. For example, as described later, the game that was developed presents standard tests of tracking ability, reaction time and short-term memory, among others. These are incorporated into the game context as piloting a shuttlecraft through a moving force field, warding off a lightning attack by Phantom Xenoid invaders, and making a situation report to the HUB Commander, respectively. The aim of these procedures was not to fool or deceive the individual, rather it was to create a context in which the person is absorbed by the activity, forgets about being assessed on some particular function, and is motivated to perform to the best of his or her ability.

Third, the task must have strong motivational properties. The importance of task motivational properties became apparent in a different

study we were conducting for the Air Force at that time. The effects of sleep deprivation on performance and biochemical indicators of fatigue were evaluated, using 28 subjects undergoing up to 60 hr of sleep deprivation. The work/rest cycles were systematically varied in this study, and during rest periods subjects played computer games similar to the one developed for this project. The subjects played the games continuously for hours, becoming more and more involved. They got into competition with each other, and often formed two- or three-man teams to go on missions together, naturally dividing into navigation, weapons and command specialists. The differences in motivation, involvement and cognitive effort during game performance, compared to the same individual's performance on traditional experimental work and performance tasks, was readily apparent and served to highlight the potential utility of this type of material in studies of continuous performance and environmental stress.

It was decided that the game should be one of skill, not luck. Successful performance would depend on the ability to make complex decisions under pressure, and to carry out these decisions effectively and rapidly. The game would also provide multiple inputs and changing situations, and require the individual to make choices and assess risks and benefits. Finally, it would provide as much freedom to decide and act as is feasible within the constraints of measurement. It was believed that these characteristics would aid in maintaining motivation over long periods of time.

Fourth and above all, the game must be an experimental tool. In essence, this implies that beneath its facade of apparently random situations and unlimited freedom to act, it would actually be a highly determined and controlled experimental task. The choice to act would be a choice among a set of circumscribed operations. Similarly, the apparently random situations would be presented from a predetermined set of situations with known parameters. The optimal decision for a particular situation would be known beforehand, and individual decision adequacy would be evaluated against this optimum value. The processes an individual uses to arrive at and execute a decision would be made overt and measurable through the timing and sequence of the interrogatory, action and response commands used to play the game, in relation to the situation presented on the display screen. Thus, all measurement would occur inside the computer, and be based on the real-time actions of the individual. This approach would allow the task to remain a challenging and absorbing game to the individual, and simultaneously, an experimental tool for the evaluation of multiple cognitive performance skills for the researcher.

2. Anticipated uses of the task: The anticipated uses of the task were also clarified during proposal preparation. We envisioned the task as an experimental tool to evaluate the influence of the following types of factors on complex human abilities:

- \* Environmental stress (heat, cold, vibration, protective gear, etc.)
- \* Fatigue and recovery from fatigue (sleep deprivation, circadian rhythms, work/rest cycles, napping behavior)

- \* Drugs of abuse
- \* Individual differences (age, sex, IQ, personality factors)
- \* Sustained performance demands (workload, duty cycle)
- \* Biochemical and physiological correlates of performance under stress
- \* Identification of areas of likely human error under stress, and enhancement of performance efficiency and effectiveness
- \* Training issues and complex task learning

The variety of contemplated uses also dictated certain restrictions on the form of the task. The task should be highly automated and computerized. The programming language should be a standard one, widely available, and known to a variety of personnel. The computer equipment should also be widely available, and not include esoteric or highly specialized hardware or software routines or functions. Instructional materials and operational procedures should be easily understood and performed. Performance data should be reported out in an easily readable form. A wide variety of performance measures should be included since some measures may be more relevant to certain uses than others. Finally, the computer programs should be constructed as far as possible in modular form to facilitate program changes necessary to meet additional specific research needs.

#### B. Project Activities

Initial project activities focused on evaluation of existing computer war games. Five existing computer war games were identified and located. Detailed information, paper documentation and/or computer disk copies of the source programs for these versions were obtained. Four of the five versions were made operational on our computer system. These versions were compared and evaluated. Details of the internal program architecture, the mathematics involved in program parameter setting, and the logic underlying operator command/control procedures were derived and flow-charted.

This review indicated that for the majority of these games, many of the command and control options were more "window dressing" than functional options. The random number generator was used extensively to put a chance factor into operator calculations and game operations. When choices were offered, as for example between using one weapon system over another, often no means was provided for the operator to make a considered judgement. An important observation, however, was that the games were extremely motivating and absorbing to the player. Individuals would literally spend hours playing the games, calculating strategic and tactical moves, and comparing their performance against the performance of others. In short, the games

had an absorbing, challenging quality that personally involved the individual, and caused him or her to continue trying to perform optimally for long periods of time.

Our review also indicated that the games had basically the same scenario. One individual is placed in command of a small elite force and is sent on missions against a superior hostile force. If the captain has the "right stuff", and uses his or her brains and ability with skill and daring, he or she has a chance to win. If the mission is completed successfully, then the captain has the opportunity to go on to greater challenges. Eventually his success becomes the challenge for others to beat. The underlying psychological factor in this scenario is an extremely powerful and pervasive epic myth. It is at least as old as the Bible, and is summarized in the biblical quotation that an individual grows best by "going in harm's way" and learning to overcome adversity. This myth underlies the appeal inherent in such diverse entities as the Odyssey, Robin Hood, the WW II defense of Britain, Generals Patton and MacArthur, and the recruiting slogan "... a few good men". The myth has the capacity to tap directly into feelings of, and desires for, innate individual superiority. These feelings and desires exist in many people, and can be used to motivate high levels of performance over extended periods of time. Recognition of these factors led to the decision to incorporate the cognitive task to be developed inside a scenario which emphasized this underlying epic myth.

The task developed in this project is called the Strategic and Tactical Assessment Record (STAR). Developmental activities were conducted between May 1, 1980 and April 30, 1981. In addition to the project activities described above, staff created the game scenario, decided on the type and function of each game element, and devoted considerable time to operationally defining the measures to be obtained. A comprehensive plan for the construction of STAR was prepared and programming, testing and debugging activities conducted. The end result of these activities was an initial Report of Deliverables. This report, together with supporting documents, was submitted to the project COTR on July 16, 1981, in compliance with page 2, section A, paragraph 3 of Contract DAMD17-80-C-0075.

#### C. Description of the STAR (Strategic and Tactical Assessment Record)

The purpose of this section is to provide a general description of STAR, its operating environment, and its performance requirements.

1. Overview: The task is controlled and operated through use of a set of five computer programs. These are: GALAXY, PATH, STAR, SHORT, and OUTPUT. The general procedure is for the researcher to use GALAXY prior to an experiment to set up all the parameters of a mission, or of a series of missions. Thus, GALAXY provides the major means of experimental control over the task. It will determine the distribution and concentration of enemy forces, the difficulty level of the game, mission time, and the resources available to the captain (player).

The PATH program is used only during training. It is an automated, visual motor tracking-teaching program. Once the researcher sets in the initial performance progression criteria, PATH will interact with the trainee to present a series of training trials. PATH collects performance data, provides trial-by-trial performance feedback, and if criteria are met, bells go off, the trainee is congratulated and the program automatically changes to the next more difficult criterion. It stops when the trainee meets the tracking difficulty level criterion used in STAR.

STAR is the program that actually presents the game. It requires the researcher to set in initial subject and data file identification information. Beyond that, STAR interacts totally with the captain. Raw data are collected on the timing and sequence of the commands issued by the captain, and these data are stored in categories related to particular situations encountered during a mission.

SHORT is the program that is automatically called after each mission. Its purpose is to present a Mission Debriefing Report to the captain. It tells the captain how well he or she performed in four main areas of command skills, and provides information on four areas where error correction is required. It also presents a summary of mission effectiveness, and tells the captain if performance will result in promotion, demotion or remaining at present rank. Finally, it shows the criterion to be met to obtain the next level of command rank, and it provides a few choice words if the captain made major command errors.

The final program is OUTPUT. This program is used by the researcher to reduce the raw performance data collected by STAR into all the measurement variables, or combinations of variables, actually collected for analysis purposes. Data can be printed out on paper record or presented on the computer terminal. The output program collects selected measures of human performance and subjective state obtained during a mission. The program is designed to present 80 performance measures divided into 8 major areas of human function. Table 2 in Section III-C-4 presents the measures obtained. An additional program option can be used to present all of the measures derived from the STAR task,

The attachments to this report present complete documentation of the above computer programs. This documentation includes: (a) computer listings of the source programs for each component of the STAR task; (b) copies of each source and run program on 8-in. floppy disk; (c) operational and descriptive material necessary to run the programs; and (d) sample printouts of the mission debriefing report, selected output, and total output programs.

2. Operating environment: The STAR programs are configured to function in the following computer environment:

Computers: DEC 11/03, 11/23, 11/34

Operating system: DEC RT-11 F/B, version 4

Program language: DEC FORTRAN IV version 2.5

Hardware peripherals: Programmable real-time clock,  
Flexible or hard disk

Terminal: DECSCOPE VT-52 or VT-100 series using VT-52  
cursor control codes (4800 baud rate)

Library calls: Loading and inspection of addresses. Timing  
interrupt completion routines from pro-  
grammable clock

Over the course of the project, various configurations of equipment were evaluated for ease of use, communication flow, and performance monitoring effectiveness. The environment described below represents an efficient, cost-effective approach, since it allows two subjects to be run simultaneously by one researcher. This environment consists of a central monitoring station equipped with two, 12-in., closed-circuit, black and white television monitors, and two audio intercom masters. Using this equipment the researcher can continuously monitor the game performance of, and interact with, two subjects performing STAR simultaneously in separate rooms. The performance rooms are each equipped with a DEC 11/23 computer, a DEC VT-100 series terminal display, and an open audio intercom. Each TV monitor is "slaved" to a VT-100 display such that the researcher can follow the real-time action of a mission in each performance room. If the subject speaks or has a question, he or she can be heard over the open intercom and a response made if appropriate.

3. Task scenario: This section will describe the overall context of STAR. Detailed information on the operation of command and control functions, information display characteristics, and performance options is presented in the Training Manual included in the attachments to this report.

The task is embedded in the context of a futuristic war where the operator is required to continuously make strategic and tactical decisions in order to accomplish the mission. At the disposal of the operator is an array of sophisticated battlefield control systems. These systems are activated through operator interactions with the computer terminal. The results of each command decision are displayed to the operator on the computer terminal. The primary measures of performance and decision adequacy are derived from the timing and type of interrogatory, response and action commands issued by the operator during the mission.

The operator assumes the role of captain of a Federation cruiser (VENTURE) and is sent repeatedly on missions which can vary in difficulty and duration. The task on any particular mission is to locate and destroy a specified number of alien cruisers (Xenoids) within the time limits of that mission, and in the most energy efficient fashion possible.

The successful accomplishment of a mission is dependent on the knowledge and skill of the captain in using the automated battlefield control systems available onboard VENTURE. Available systems include: short- and long-range scanners to help locate the Xenoids and navigate through the galaxy; two types of offensive weapons (phasers and photon torpedoes) for use during attack; defensive energy shields to protect the VENTURE from the effects of attack; navigation systems for movement over short distances and for long-distance travel; and an on-board computer system that (1) provides information on the extent of damage and the status of various system energy levels, (2) presents a visual long-term memory display of all previous long-range scans requested during a mission, and (3) allows the captain to re-allocate energy resources from the onboard reserve supply to the various systems of the cruiser.

On each mission, the captain is required to balance risks and assess benefits in accomplishing the assigned task. The captain starts each mission without adequate fuel or armament to complete the mission. These items can be replenished by docking at the Home Universe Base (HUB), the command training cruiser in orbit above his galaxy. The docking procedure incorporates a standardized tracking task. The captain must calculate risk/benefit ratios during attack situation, in plotting navigation courses, and when using the defensive shields, all of which drain the onboard energy supply. In addition, the captain has only a finite period of time to complete the mission. All operations of VENTURE both subtract time from the allowable total for the mission and require energy. Thus, the captain cannot simply travel blindly over the galaxy; rather, he or she must develop an efficient search strategy that balances mission goals against energy, armament supplies and time. Similarly, enemy hits can make certain systems of the VENTURE inoperable. The captain then has to make decisions between effecting quick repairs, which cost energy units in proportion to the type and extent of damage, or returning to the HUB which costs mission time.

The captain is also faced with additional hazards. Present in the galaxy during each mission are variable numbers of more sophisticated enemy vessels. Phantom Xenoids are not detectable by VENTURE scanners. They appear suddenly and unexpectedly, and unless the captain takes immediate defensive action, can inflict serious damage. Thus, the captain must exercise continuous vigilance to avoid damage to VENTURE. A second type of more sophisticated enemy vessel is the Super Xenoid. This type of enemy is detectable by VENTURE scanners; however, it is camouflaged so that it appears to be a typical Xenoid cruiser. Only when the captain enters a quadrant containing a Super Xenoid does its presence become known. Super Xenoids have the technical capability to drain the life support system of the VENTURE at a fast steady rate, and begin to do so as soon as VENTURE enters the quadrant. The captain must then decide how to destroy the Super Xenoid with minimal energy loss. He or she also has to monitor and maintain the life support systems at an appropriate level, or risk having the cruiser destroyed.

The captain begins each mission by flying a shuttlecraft out to the HUB to take command of a vessel. VENTURE then enters the galaxy at

quadrant 1,1 (the galaxy is divided into 64 quadrants arranged as an 8 x 8 grid, and each quadrant is in turn further divided into 64 sectors on the same basis). The location of Xenoids and stars is unknown to the captain. The captain issues movement commands via the terminal, and the VENTURE travels from quadrant to quadrant searching for the Xenoids. The captain uses short-range scans to examine the location of contents of the quadrant in which the VENTURE is presently located. If Xenoids are present, the captain adjusts the shields accordingly and the Xenoids are attacked using either torpedoes or phasers. The results of these actions are continuously displayed on the terminal. VENTURE can be destroyed if a Xenoid attack gets through the energy shields of the cruiser, the life support system is drained, or if faulty navigation results in a collision. Once the action is completed within a particular quadrant, the captain resumes the search. Generally, a long-range scan is requested to reveal the number of Xenoids and stars within the surrounding eight quadrants, but not their exact locations. The captain then plots a course, checking for possible collisions, and issues a navigation command. Once at the desired new location, he or she attacks the Xenoids present. At other times, the captain will move to locations in the galaxy which are outside the previous range of the long-range scan and begin a search of the new area.

Prior to any movement of the VENTURE, the captain must use the terminal to enter into the mission log the purpose of the movement (search, attack, maneuver, evade) and the desired destination (e.g., quadrant 4,6). At the beginning and end of each mission, the captain makes a status report to the HUB commander. During the mission, each time the captain wishes to return to the HUB he or she must state the specific reason and provide a status report. Thus, the use of VENTURE battlefield control systems incorporates unobtrusive measures of the captain's intent and reasons for actions. This, and other features of the program, allow analysis of errors and sequences of activities that led to error.

The captain continues making between-quadrant strategy decisions and within-quadrant tactical decisions until the mission goal is accomplished or the mission duration is met. The captain is then sent on another mission, and the above sequence repeats. Table 1 presents the 11 major command options available to the captain. Figure 1 presents the basic information display used by the captain during the mission.

TABLE 1

COMMAND AND CONTROL SYSTEMS OF THE VENTURE

<u>System Type</u>	<u>Computer Keyboard Activation Key</u>
Navigation (NAVCON)	Period (.)
Short-range Sensor Scan (SRS)	One (1)
Long-range Sensor Scan (LRS)	Two (2)
LRS History Scan	Three (3)
Photon Torpedo Control	Four (4)
Phantom XENOID Defense Shield	Five (5)
Phaser Control	Six (6)
Resource Allocation Command	Seven (7)
Docking Control	Nine (9)
Cancel Command	Zero (0)
Surrender	L (Alpha L)

Command ?

\_\_\_\_\_ Message Center \_\_\_\_\_

\_\_\_\_\_ Status Report \_\_\_\_\_

1 Shields =

2 Torpedoes =

3 Phasers =

4 Fuel =

5 Life Support =

6 Repairs =

7 Reserve =

Present Energy Level  
( ) ← Torpedoes

Name of Scan Selected

1									
2									
3									
4									
5									
6									
7									
8									

8 x 8 Grid  
Used to Present  
Short Range Scan  
Long Range Scan  
Long Range Scan History

Type of Scan Presented  
Changes on Captain's  
Command

Mission Time Remaining

Present Quadrant Position

Number of Xenoids to be Destroyed

Figure 1 -- Basic Information Display Pattern of the VENTURE Cruiser

4. Performance measures: The apparently random situations encountered by the captain, and the apparently unlimited freedom of action, are actually a carefully constructed facade. The dispersion of enemy cruisers and their relative concentrations are mathematically determined by the GALAXY program beforehand. Similarly, the distribution and density of stars are also predetermined. These factors are used to alter task difficulty level. The freedom to act is in reality a freedom to choose from a predetermined set of operations. In any attack situation there is a mathematically determined optimal decision, given the position of the VENTURE, and the position of the enemy. The adequacy and speed of the captain's decision making are evaluated against this optimum. Similarly, search strategy decisions and information resource usage are evaluated against additional known parameters of enemy dispersion and placement. Thus, this complex task provides a number of measurement parameters that can be operationally defined and obtained unobtrusively without interfering with ongoing activity.

It is anticipated that STAR will be used to examine higher order functioning under long-term continuous performance conditions as well as to assess effects of various types of environmental stressors. Since the research paradigm is new, a number of measurement parameters have been included in the design of the task. Some of these measures will undoubtedly prove to be more useful under specific conditions than others. Table 2 presents a listing of major measures obtained during each mission. These measurement parameters are calculated by the OUTPUT program, and represent a selection of the total measures available from each mission. A second option of OUTPUT can also be used to present all currently available measures.

#### IV. PHASE II: TRAINING AND TASK EVALUATION

At the conclusion of Phase I, continuation contracts were awarded for the period May 1, 1981 through September 30, 1982. The goals for the second phase of project activities were to: (1) develop a training protocol and training materials for STAR, and test these on a sample of 12 human volunteers (6 men and 6 women); (2) conduct a study to evaluate the test-retest reliability of STAR performance measures; (3) conduct a study to establish and test three levels of task difficulty; (4) conduct a study to evaluate the effects of time pressure on performance; (5) conduct a study to evaluate STAR performance under sustained operations conditions; and (6) prepare all materials necessary for an operational human performance testing package. All of the above goals were accomplished, with the exception of No. 5 above. After discussion with the COTR, it was decided that due to funding constraints, the remaining project funds would be used to document the findings generated to date, and to fully accomplish goal No. 6 above.

The activities involved in relation to each of the above goals are described in detail in the following sections of this document. In all, a total of 22 human volunteers who met U.S. Army enlistment standards participated in this phase of the program. Each received a full and complete explanation of the activities required, and each signed and received a copy of the Informed Consent Statement shown in Figure 2. Each subject was paid for his or her participation at the rate of \$3.50 per hr.

TABLE 2

SAMPLE OUTPUT PRINT SHOWING MAJOR PERFORMANCE PARAMETERS OF STAR

DATA FILENAME : RK1:H06B01  
 DATE OF MISSION : 3/25/82  
 START TIME : 09:24:49  
 COMMANDER WAS CD  
 GALAXY ENCOUNTERED : 00921

OVERALL PERFORMANCE EFFICIENCY

MEAN TIME PER ENEMY CRUISER DESTROYED (SEC)	= 52.10
MEAN ENERGY UNITS REQUIRED PER ENEMY DESTROYED	= 310.00
NUMBER OF FRIENDLY CRUISERS LOST	= 1
NUMBER OF ENEMY CRUISERS DESTROYED	= 20

PSYCHOMOTOR SKILLS

## STANDARDIZED VISUAL-MOTOR TRACKING TASK

PERCENT TIME ON TARGET	= 89.9
ROOT MEAN SQUARE (RMS) TRACKING ERROR	= 2.14
MEAN ABSOLUTE ERROR	= 1.66

## STANDARDIZED REACTION-TIME TASK

MEAN REACTION TIME (MSEC)	= 675.
PERCENT SUCCESSFUL TARGET DETECTION	= 80.0
NUMBER OF ERRORS	= 1

PERCEPTUAL ACCURACY AND SPEED

## TARGET AIMING

NUMBER OF TORPEDOES FIRED	= 1
PERCENT TORPEDOES ON TARGET	= 100.0
MEAN ABSOLUTE TORPEDO COURSE ERROR	= 0.00
MEAN TORPEDO EXECUTION TIME (SEC)	= 25.090

## NAVIGATION

NUMBER OF QUADRANT ENTRIES	= 22
PERCENT QUADRANT ENTRIES AS INTENDED	= 95.5
MEAN NAVIGATION EXECUTION TIME (SEC)	= 7.461

MEMORY FUNCTION

## STANDARDIZED MEMORY RECALL TASK

## OVERALL PERFORMANCE ACCURACY AND SPEED

PERCENT CORRECT RECALL	= 75.8
MEAN RESPONSE TIME (MSEC)	= 1817.

## RECENT MEMORY (5 TEST ITEMS)

PERCENT CORRECT RECALL	= 80.0
------------------------	--------

TABLE 2 (continued)

MEAN RESPONSE TIME (MSEC)	=	2061.
MEAN POST-PRE MISSION RECALL RESPONSE TIME (MSEC)	=	428.

## LONG TERM MEMORY (2 TEST ITEMS)

PERCENT CORRECT RECALL	=	62.5
MEAN RESPONSE TIME (MSEC)	=	1054.

## ERRORS OF MEMORY STORAGE DURING COMMAND OPERATIONS

NUMBER OF LONG RANGE SCANS WITH NO NEW INFORMATION	=	6
NUMBER OF LONG RANGE SCAN HISTORIES WITH NO NEW INFORMATION	=	11
PERCENT OF WEAPON AND NAVIGATION COMMANDS ISSUED WITHOUT ADEQUATE RESOURCES AVAILABLE	=	3.2

# INFORMATION PROCESSING

---

## STRATEGIC

---

## INFORMATION RESOURCE USAGE

## AMOUNT

PERCENT OF TOTAL COMMANDS USED FOR:	
LONG RANGE SCANS (LRS)	= 14.5
LRS HISTORY SCANS (LRSH)	= 16.9
COMBINED LRS AND LRSH	= 31.3
NUMBER OF LRS COMMANDS	= 24
NUMBER OF LRSH COMMANDS	= 28

## FREQUENCY

MEAN TIME BETWEEN LRS COMMANDS (SEC)	=	39.790
MEAN TIME BETWEEN LRSH COMMANDS (SEC)	=	30.602

## DURATION

MEAN TIME TO EXTRACT INFORMATION PER LRS COMMAND (SEC)	=	2.372
--	---	-------

## EFFICIENCY

PERCENT REDUNDANCY IN LRS INFORMATION REQUESTED	=	65.8
PERCENT NONCONTIGUITY OF QUADRANT MAPPING IN LRS INFORMATION	=	16.7
MEAN NUMBER QUADRANTS OF NEW INFORMATION PER LRS COMMAND	=	2.7
NUMBER OF NAVIGATION COMMANDS ISSUED FOR PURPOSES OF SEARCH	=	12
UNITS OF FUEL EXPENDED IN SEARCHING FOR ENEMY	=	865

## TACTICAL

TABLE 2 (continued)

### ACCURACY

MEAN PHASER PAYLOAD CALCULATION ERROR	=	-7.
PERCENT SUCCESSFUL PHASER COMMANDS	=	85.7

### DURATION

MEAN TIME TO CALCULATE PHASER PAYLOAD FOR ONE ENEMY CRUISER (SEC)	=	24.934
MEAN TIME TO CALCULATE PAYLOAD FOR TWO ENEMY CRUISERS (SEC)	=	10.015
MEAN TIME TO CALCULATE PAYLOAD FOR THREE ENEMY CRUISERS (SEC)	=	13.595
COMBINED MEAN PAYLOAD CALCULATION TIME WHEN STRESSOR (SUPER XENOID CRUISER) IS PRESENT (SEC)	=	9.014
COMBINED MEAN PAYLOAD CALCULATION TIME WHEN STRESSOR IS NOT PRESENT (SEC)	=	19.099

### EFFICIENCY

MEAN ENEMY DESTRUCTION TIME OVER QUADRANTS (SEC)	=	31.294
PERCENT OF ALL COMMANDS USED SPECIFICALLY FOR NAVIGATE ATTACK	=	6.6
AMOUNT OF FUEL USED SPECIFICALLY FOR ATTACK	=	700
NUMBER OF NAVIGATE ATTACK COMMANDS	=	11

### DECISION MAKING

#### FUNCTIONAL COMMAND BREAKDOWN

NUMBER OF COMMANDS ISSUED PER MISSION	=	166
PERCENT OF TOTAL COMMANDS USED FOR NAVIGATION	=	28.3
PERCENT OF TOTAL COMMANDS USED FOR INFORMATION SEEKING	=	31.3
PERCENT OF TOTAL COMMANDS USED FOR WEAPONS	=	9.0
PERCENT OF TOTAL COMMANDS USED FOR DEFENSIVE RE-SUPPLY	=	0.0
PERCENT OF TOTAL COMMANDS USED FOR OFFENSIVE RE-SUPPLY	=	5.4
PERCENT OF TOTAL COMMANDS USED FOR DAMAGE CONTROL	=	0.0

#### DECISION MAKING EFFICIENCY

OVERALL MEAN TIME PER COMMAND (SEC)	=	6.507
NUMBER OF NAVIGATIONS TO A DIFFERENT BATTLE POSITION DURING ATTACK	=	4
NUMBER OF NAVIGATIONS TO A BETTER BATTLE POSITION DURING ATTACK	=	3
PERCENT CORRECT WEAPON CHOICE (COMPARED TO A MATHEMATICAL OPTIMUM DURING EACH ATTACK SITUATION)	=	100.0

### RISK TAKING BEHAVIOR

MEAN ENERGY LEVEL IN PROTECTIVE SHIELDS ON ENTERING ENEMY QUADRANTS	=	435.8
MEAN SHIELD LEVEL DURING WEAPON COMMANDS	=	591.3
PERCENT WEAPON COMMANDS WITH INADEQUATE SHIELD PROTECTION	=	0.0

TABLE 2 (concluded)

MEAN DIFFERENCE BETWEEN ACTUAL SHIELD LEVEL AND MINIMUM PROTECTION LEVEL REQUIRED DURING ATTACK	=	409.1
MEAN ENERGY LEVEL TO WHICH THE CAPTAIN LETS THE LIFE SUPPORT SYSTEM DROP BEFORE ISSUING A RESUPPLY COMMAND (-1 = NO DATA)	=	-1
MEAN AMOUNT OF ENERGY ALLOCATED TO LIFE SUPPORT (-1=ND)	=	-1
LEVEL OF ONBOARD RESERVE ENERGY LEFT WHEN VOLUNTARY DOCKING COMMAND ISSUED (-1 = NO DATA)	=	0.
MEAN ENERGY LEVEL REMAINING IN SHIELDS WHEN VOLUNTARY DOCKING COMMAND ISSUED (-1 = NO DATA)	=	478.

SUBJECTIVE RATINGS (1-10 SCALE)

-----

		DOCKING:	PRE	POST
PSYCHOLOGICAL STATUS	STRESS	=	5	5
	CONFIDENCE	=	5	1
	EFFICIENCY	=	5	1
	WORKLOAD	=	0	9
PHYSICAL STATUS	FATIGUE	=	5	5

Figure 2

VOLUNTEER'S INFORMED CONSENT

I, \_\_\_\_\_  
residing at \_\_\_\_\_

hereby acknowledge and certify to the following:

1. I hereby volunteer and consent to participate as a subject in a research study entitled "Fragmented Sleep and Cognitive Efficiency Under Stress" to be conducted at Midwest Research Institute. This study is funded by the U.S. Army Medical Research and Development Command. I understand that the purpose of this study is to develop a computer game which will evaluate human information processing and decision making abilities. I understand that I will be asked to come to MRI for several sessions of 2-3 hr in length. Prior to each session, the specific activities planned will be explained to me, and I will have the opportunity to freely choose to participate or not to participate. I understand that I will be asked to learn to play the game to the best of my ability, that I will be asked questions about my mood, my perception and my performance, and that I may be asked to play the game for extended periods of time (12 to 18 hr). In these extended sessions, I understand that my temperature will be taken and that I will be asked to perform a computerized tracking task in addition to playing the computer game. I further understand that there are no known risks associated with my participation. Finally, I understand that this is a basic research study and it is not designed to personally benefit me except to the extent that I will be contributing to scientific knowledge, and I will be paid \$3.50 for each hour of participation.

2. I have been given, in my opinion, an adequate explanation of the nature, duration and purpose of the experiment, the means by which the experiment will be conducted and any possible inconveniences, hazards, discomfort, risks and adverse effects on health that could result from my participation;

3. I understand any questions concerning procedures that might affect me or my rights as a subject will be answered fully and promptly by the project staff or by Dr. Charles Graham the Principal Investigator (phone number 816-753-7600);

4. I understand that I have the right to withdraw my consent and to discontinue participation in this experiment at any time without prejudice regardless of the status of the experiment and regardless of the effect of such withdrawal on the objectives and results of the experiment; and I also understand that my participation in the experiment may be terminated at any time by the investigator in charge of the project;

5. I understand and agree that I will inform the investigator in charge of information regarding any medications I have taken and any medical or dental care or treatment I have received 24 hours prior to the experimental session;

6. I hereby certify that the medical history information which I provide to Midwest Research Institute is correct to the best of my knowledge, and that I have informed Midwest Research Institute staff of all serious or chronic medical problems which I now have or have had in the past;

7. If I am physically injured as a result of participation in this study, emergency treatment will be provided in accordance with the policy of MRI; for additional information regarding this policy, contact the MRI Personnel Department (816-753-7600);

8. I agree that any information obtained from me by Midwest Research Institute or by its authorized representatives in connection with these experiments may be utilized by Midwest Research Institute in publications and reports without identifying me;

9. I attained the age of \_\_\_\_\_ years on my last birthday, which was \_\_\_\_\_, and that I am executing this Volunteer's Informed Consent as my free act and deed.

Executed this \_\_\_\_\_ day of \_\_\_\_\_, 19 \_\_\_\_\_

Executed in my presence and in the presence of each other

\_\_\_\_\_  
Signature of Witness

\_\_\_\_\_  
Signature of Volunteer

\_\_\_\_\_  
Title or Position of Witness

\_\_\_\_\_  
Signature of Experimenter

\_\_\_\_\_  
Residence Address

## A. Development and Testing of a Training Protocol

The specific aims of this task were to: (a) develop the necessary materials and an efficient protocol for the training of individuals to perform the STAR task; and (b) test the training procedures on a sample of 12 subjects to establish performance criteria. The final version of the training manual and task exercise problems to be used in teaching individuals to perform STAR, as well as an 8-in. floppy disk containing one demonstration mission and 30 training missions arranged in order of increasing task difficulty, are included in the attachments to this report. These missions can be used to implement the training protocol, which is described in detail in Appendix A.

1. Overview: STAR is both a new and complex task; thus, at the beginning of this project phase we did not know exactly what people would have to learn in order to perform it well, nor the level of performance they could attain given intensive practice at the task. In order to gain insight into these aspects of task performance, informal pilot studies were conducted using Institute staff. These observations indicated that the best approach to teaching task performance was to prepare a training manual that could be used to inform subjects during the first training session about the command sequences and display systems used, and the types of strategic and tactical decisions required. We also determined that subjects could benefit from working paper and pencil training exercise problems prior to actually performing the task. These exercise problems proved valuable in highlighting areas of task performance that were not thoroughly understood.

Following the pilot work, a draft training manual, a set of exercise problems, and a training protocol were developed. Four subjects were then trained to perform STAR using these procedures. Experimenter observations, subject comments and performance data were integrated and assessed. The procedures were revised, and then tested using a second group of four subjects. Necessary revisions were again made, and tested on a final group of subjects. The protocol presented in Appendix A and the material presented in the attachments is the final version of the training procedures to be used in teaching individuals to perform STAR.

2. Summary of training protocol: Briefly, the training protocol is composed of a sequence of three types of training sessions. The first session is used to introduce multiple subjects to the task. They read and receive a copy of the training manual, see a demonstration of STAR, and work paper and pencil exercise problems concerned with various command activities. The next type of session involves one-on-one coaching of the subject during performance of several practice missions. The final sequence of sessions involves the subject performing missions alone. During these sessions, the experimenter monitors real-time mission activity from another area of the laboratory, and the subjects' advancement is based solely on established automated performance criteria. These criteria, as well as a report of the strengths and weaknesses of the individual's performance in a particular mission, are presented to the subject immediately after each mission is completed.

3. Training performance criteria and measures: The first group of 4 subjects (2 male and 2 female) participated primarily as an observation group to determine how people learn to perform this complex task. A set period of 18 hr of training was established. Subject started training at the level of VENTURE Captain 15, and promotion to the next successively higher command level was based on a single criterion (destroy all Xenoids within the mission time limit, and do not lose your own cruiser in the process). No control room monitors were available during the training of this group, and coaching activity centered on the Mission Debriefing Report, subject questions and experimenter observations.

This group provided important information for the development of the protocol. Based on their reports of muscle fatigue, the training time devoted to the tracking task was reduced from the original duration of 40 min to a duration that did not produce fatigue (20 min). The training manual was changed to eliminate ambiguities, and a set of 40 training exercise problems was added to the manual. Finally, the Mission Debriefing Report format was changed to provide more specific information to subjects.

This group performed a total of 81 missions. Two subjects achieved the highest command rank (Level 1) in the training time allowed. One male required 18 missions to reach Level 1, and 1 female required 27 missions. Of the other 2 subjects, one female reached Level 12 after 15 missions, and the male reached Level 6 after 20 missions.

A second group of 4 subjects (2 male and 2 female) then entered the program, and performed a total of 151 missions under the revised protocol. All subjects achieved Level 1 status. The  $\bar{X}$  number of missions required to reach Level 1 was 22.7 (range 16 to 31 missions). The protocol was revised again, and a third group (1 male and 2 females) performed a total of 66 missions under the revised protocol. The male reached Level 1 after 24 missions, and one female reached Level 7 (19 missions). The other female subject had to discontinue participation due to schoolwork demands.

Data generated over the 288 missions performed by these subjects were analyzed. One goal of these analyses was to establish empirical time and energy criteria appropriate for each command rank. Up to this point, subjects were basically functioning under a mission time criteria, and although they did demonstrate some energy reduction, this was not an explicit criterion for promotion. Table 3 presents the time and energy criteria developed from the analyses conducted.

The establishment of these criteria, and examination of other STAR measures also allowed us to change a number of the game parameters so as to bring them into agreement with the actual performance values and practices observed. For example, the reaction time task incorporated into STAR turned out to be too easy, and consequently the time allowed was cut in half. Similarly, subjects performed the tracking task much better than we had expected. Performance criteria were raised for this task. Subjects had also learned various "tricks" to circumvent the intent of the program. For example, several had discovered that once the program set up the positions of

TABLE 3

PROMOTION CRITERIA FOR VENTURE CAPTAIN LEVELS 10 THROUGH 1

<u>Rank</u>	<u>Time</u> <u>(sec/Xenoid)</u>	<u>Energy</u> <u>(units/Xenoid)</u>
10	-	-
9	Destroy all Xenoids without destroying VENTURE	
8	88	330
7	80	315
6	73	300
5	66	285
4	60	270
3	55	255
2	50	240
1	45	225

enemy cruisers in the galaxy, these positions remained fixed. These subjects then developed the technique of moving just inside an enemy quadrant and determining the best position for an attack. They would then leave and re-enter the quadrant at that position, and destroy the enemy with remarkable speed and efficiency. The program was changed such that the positions of enemy cruisers automatically rearranged on re-entry into a quadrant. The analyses also allowed us to set onboard energy availability parameters into the game that were more realistic in light of the actual energy usage of the subjects. Finally, we were able to set an initial cutoff point in the training protocol for discontinuing a subject (make first promotion within 8 missions).

Following these changes, an additional 11 subjects (4 males and 7 females) were entered into the program, and performed a total of 533 missions on both the old and new versions of STAR. Three subjects were discontinued because they could not achieve the first promotion within 8 missions. One other subject did not continue past the first session due to illness. The remaining 7 subjects all achieved Level 1 performance criteria. The X number of missions required to reach Level 1 was 23 on the old game and 11 on the new game. Six of the 11 subjects who were trained originally were also brought back and given training on the new game. All achieved Level 1 performance criteria, within an average of 13 missions.

In summary then, a total of 22 subjects (13 women and 9 men) participated in the development of the training protocol. Fourteen subjects reached Level 1 performance criteria (7 men and 7 women). The average number of missions required was approximately 22, and the sex of the subject was irrelevant to the speed of skill acquisition. Of the remaining subjects, 3 were discontinued for training deficiencies, 2 could not reach Level 1 even with retraining, 1 became ill, and 1 discontinued due to schoolwork. Examination of subject background factors indicated that those most likely to perform well at STAR had had college level courses in one or more of the

following areas: engineering, computer science; mathematics; or the sciences.

The following figures (Figures 3 through 6) present representative acquisition data obtained from subjects who reached Level 1 during training. Figure 3 presents data concerned with acquisition of tracking skill using the PATH tracking-training program. As is common with motor skills learning, males required fewer trials to reach criterion (3 sequential 30-sec trials at  $\geq 90\%$  time in PATH) compared to females, although the male and female performance distributions overlap. Overall, training was accomplished in 62 trials (31 min of time on task). Figure 4 shows the mean mission time per Xenoid (mission time divided by number of Xenoids destroyed). As skill increased, mean time to destroy Xenoids decreased. Note that during the first mission, subjects were being coached, and typically performed better than when they worked more independently during the second mission. Figure 5 shows the gain in overall performance efficiency as a function of training, and Figure 6 shows acquisition of a specific information-gathering skill.

#### B. Establishment of Three Levels of Task Difficulty

1. Introduction: The purpose of this study was to establish three levels of task difficulty (easy, medium and hard), and to test the effects of each level on task performance. Based on previous pilot work, two factors appeared to be major determinants of task difficulty: (1) the density, or number, of stars encountered in the galaxy during a mission; and (2) the dispersion of enemy units within the galaxy.

In creating a galaxy configuration, the GALAXY program randomly assigns stars to the various quadrants beginning in quadrant 1,1 and going sequentially through to quadrant 8,8 until a standard distribution is completed. In order to increase star density, the GALAXY program was changed to include a multiplication factor. The easy difficulty level was defined for this variable as the standard star distribution. The medium level multiplied the number of stars in each quadrant by two, and the hard difficulty level multiplied star distribution by three. Thus, difficulty level was defined for this variable as either single, double or triple the number of stars normally encountered during a mission. Increased star density was expected to have a direct effect on navigation variables.

The second variable, Xenoid dispersion, is also controlled by the GALAXY program. Each of the 64 quadrants of a galaxy can contain from zero to three Xenoids, and include Phantom and Super Xenoids. The operator specifies the total number of Xenoids to be placed in a galaxy, how many one, two, and three Xenoid quadrants there will be, the number and percentage of Phantoms and Supers to be included in the Xenoid-containing quadrants, and provides a starting or entry number into the Random Number Generator. The program then uses an exhaustive random assignment procedure to place the above elements in the galaxy to be created. This process results in Xenoid distributions that differ markedly from one another.

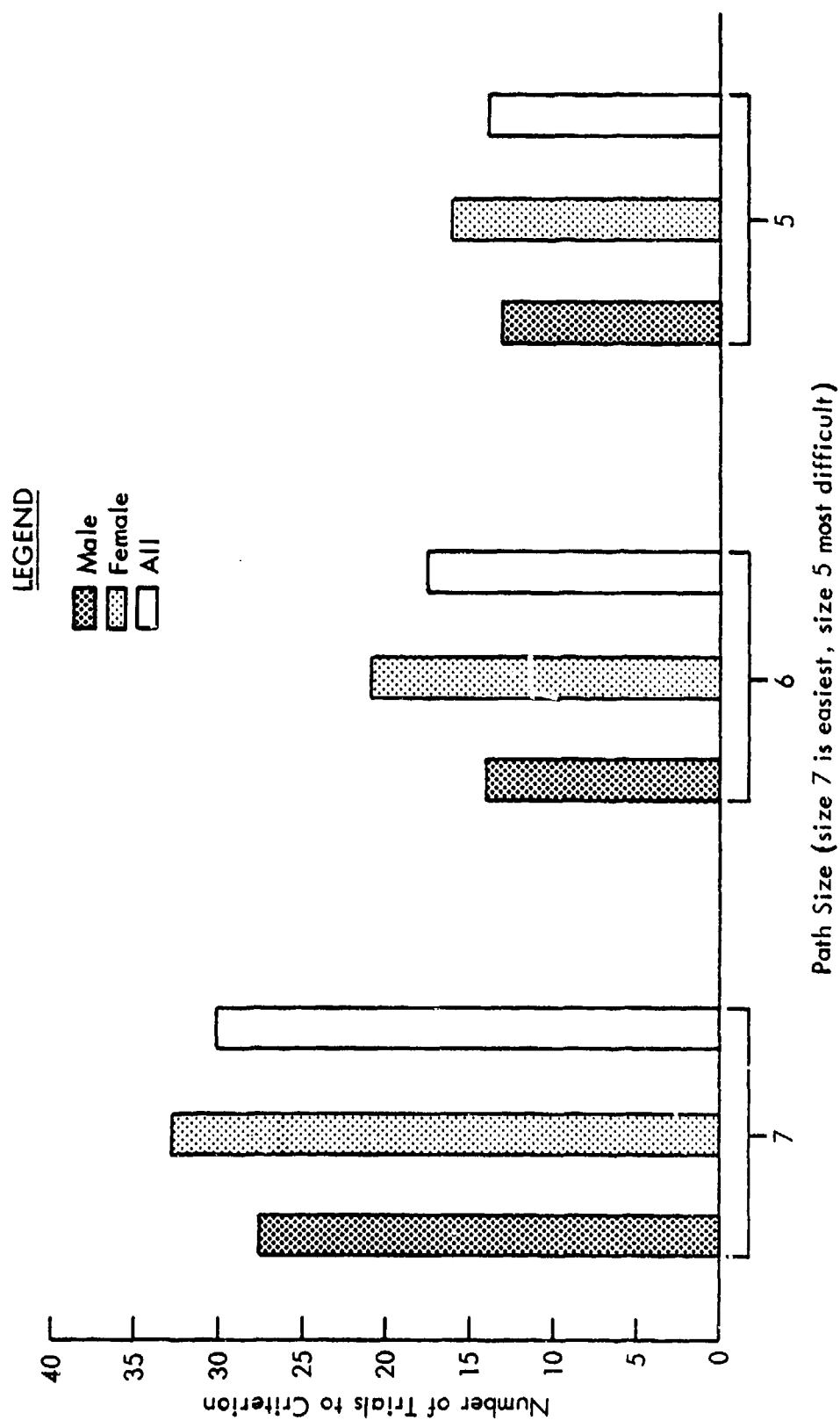


Figure 3 - Acquisition of Tracking Skills Using the PATH Program

# INITIAL TRAINING MISSIONS: MEAN TIME PER XENOID

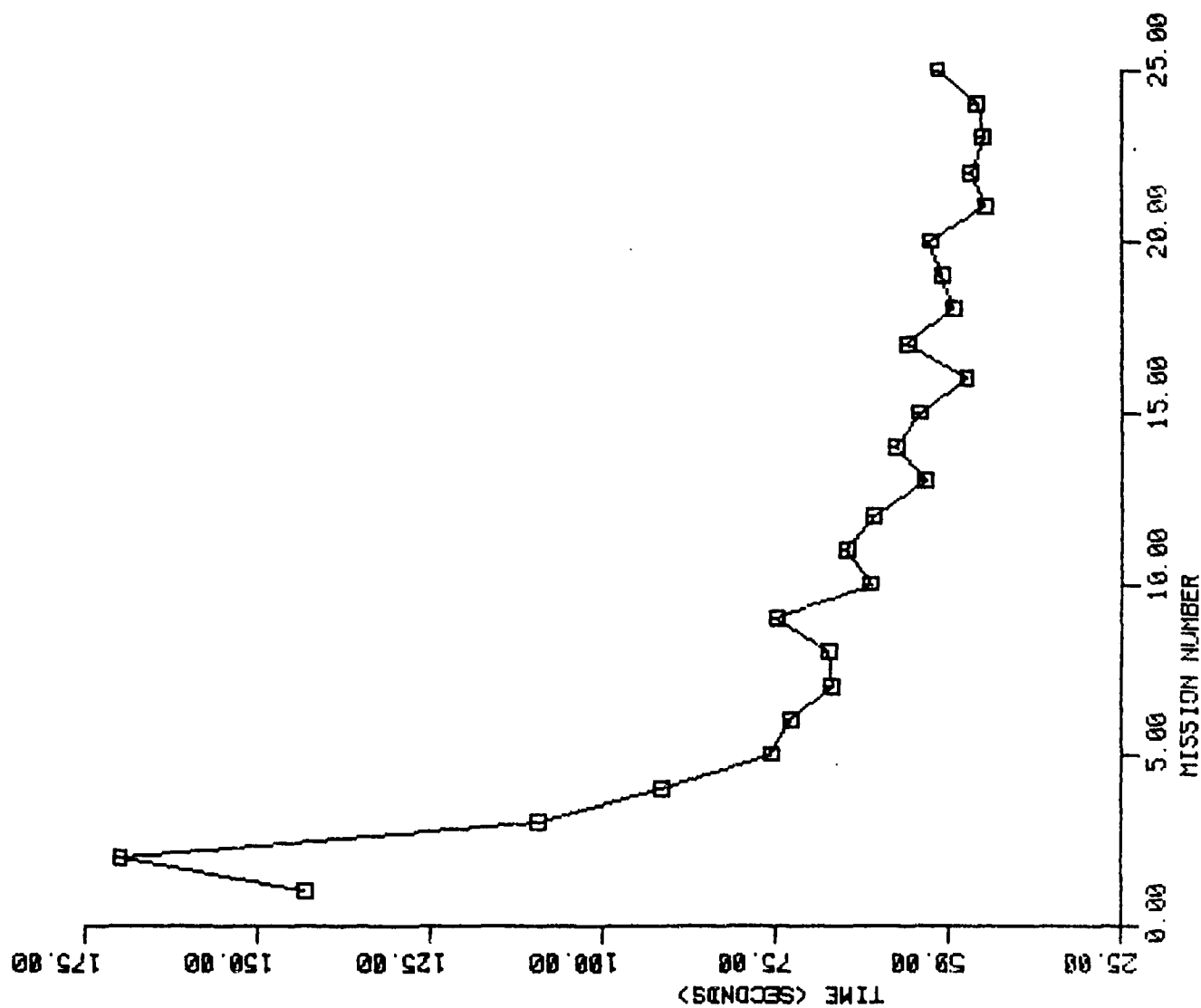


Figure 4 - Change in Mean Time Required to Destroy Xenoids as a Function of Training Trials

# INITIAL TRAINING MISSIONS: MEAN TIME PER COMMAND

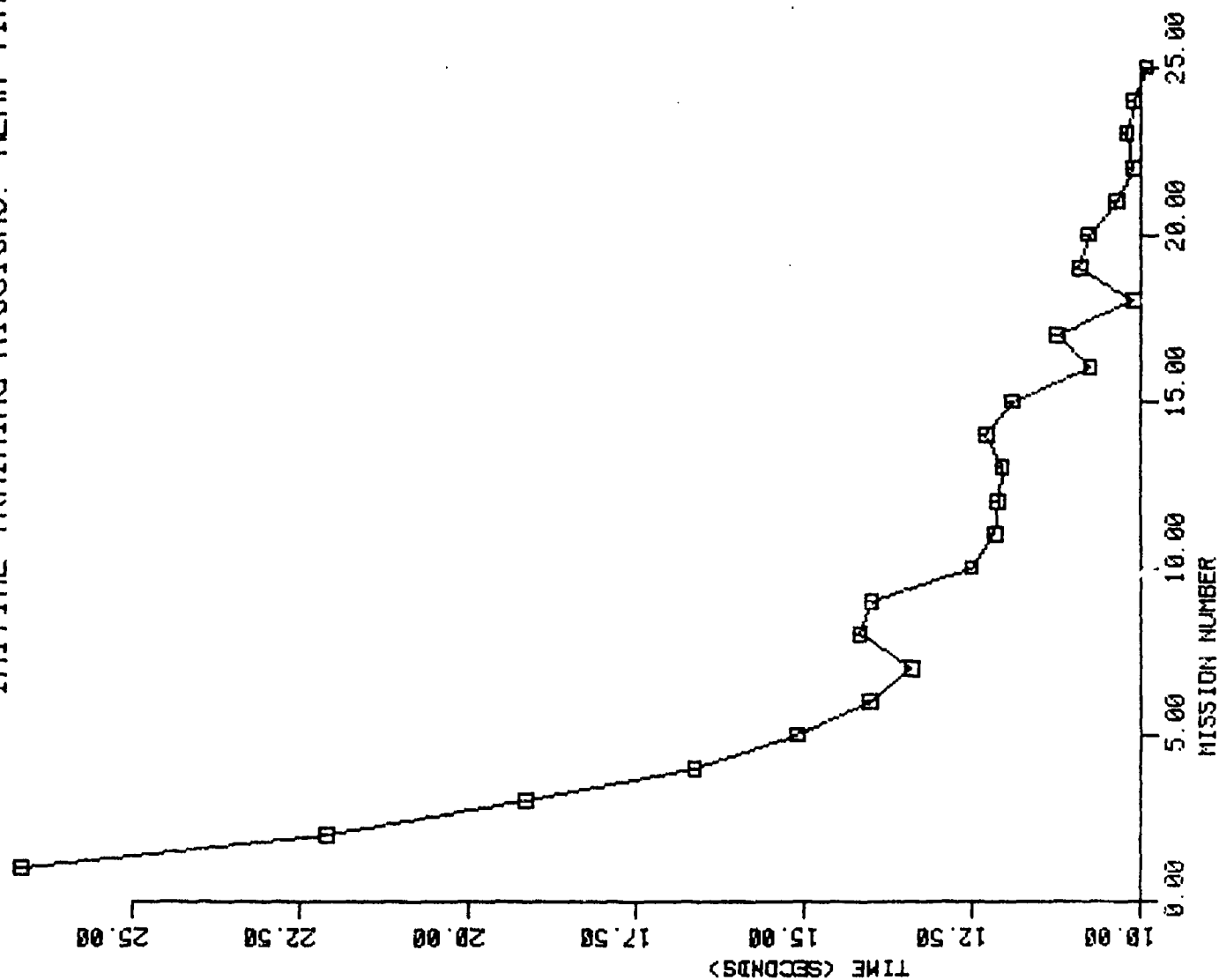


Figure 5 - Mean Time Per Command, A Measure of Overall Performance Efficiency, as a Function of Training Trials

# INITIAL TRAINING MISSIONS: MEAN TIME BETWEEN LRS'S

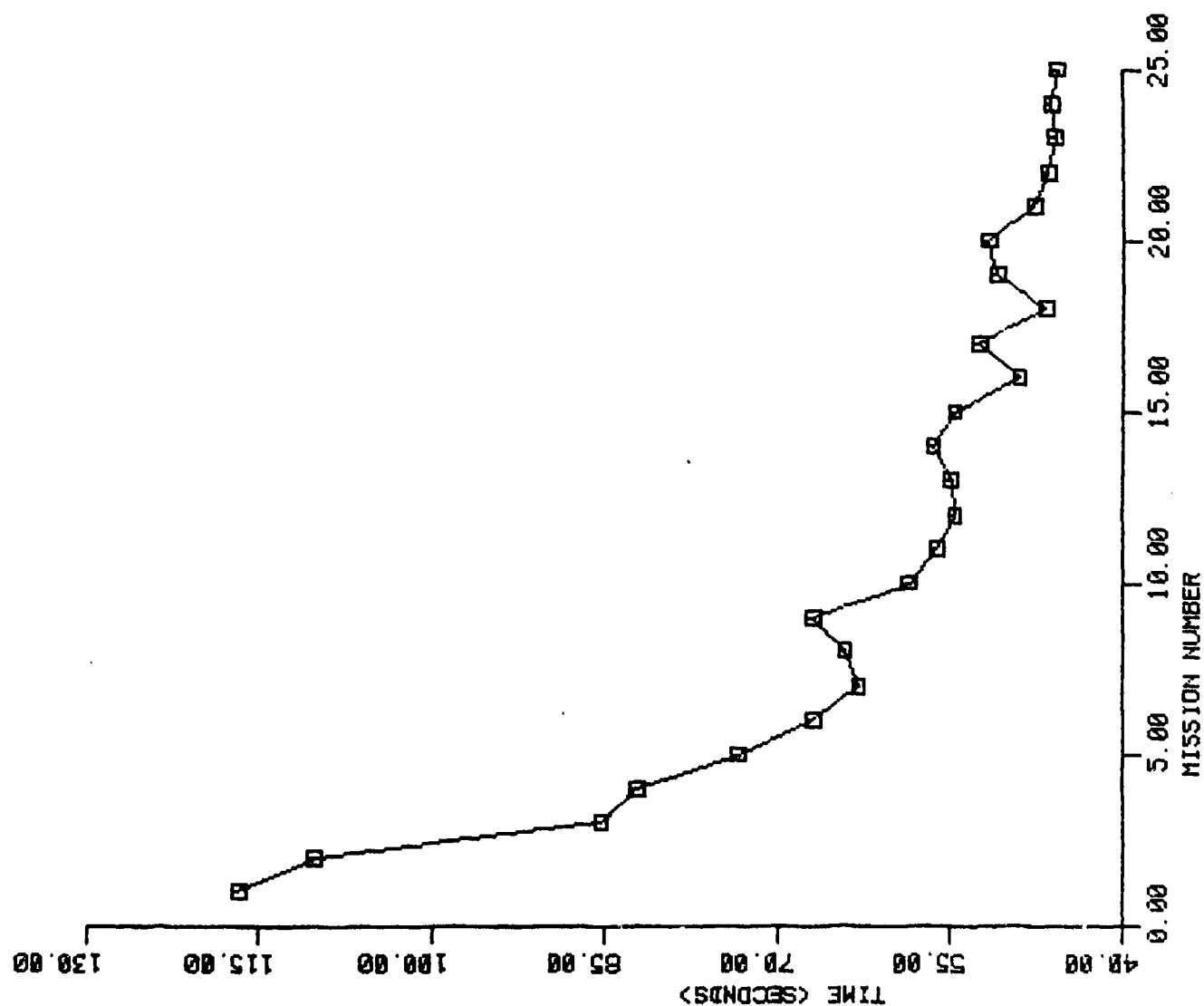


Figure 6 - Mean Time Between Long-Range Scans as a Function of Training Trials

We used a mathematical procedure to characterize and control these distributions. This procedure allowed us to vary distribution in order to alter task difficulty level. The procedure is similar to one used to characterize the dispersion of artillery hits in a target area. The 64 quadrants of a galaxy are thought of as being equivalent to the target area, and the placement of each Xenoid is thought of as equivalent to a "hit." The procedure determines the center, or mean ( $\bar{X}$ ), of the distribution, and characterizes the dispersion of Xenoids around the  $\bar{X}$ .

Maps of the 30 training galaxies were plotted showing the location of all Xenoids. The distribution in each galaxy was determined using the following procedure. Each galaxy contains eight columns and eight rows of quadrants. Each column and each row was numbered 1-8, respectively. These number designations were multiplied by the number of Xenoids in each column and each row. The totals were added together, divided by N, and the distribution mean obtained. Xenoid dispersion was calculated by obtaining the square root of the  $X^2 + Y^2$  distance from the  $\bar{X}$  to each Xenoid, and dividing the grand total by N. Over the 30 training galaxies, this procedure resulted in the distribution of Xenoid dispersion factors shown in Table 4 below.

TABLE 4  
DISPERSION OF ENEMY UNITS IN 30 TRAINING GALAXIES  
(Arranged in Order of Increasing Enemy Dispersion)

<u>Galaxy No.</u>	<u><math>\bar{X}</math> Dispersion (in quadrants)</u>	<u>Galaxy No.</u>	<u><math>\bar{X}</math> Dispersion (in quadrants)</u>
1	1.93	16	2.94
2	2.07	17	2.95
3	2.34	18	3.01
4	2.36	19	3.06
5	2.41	20	3.08
6	2.53	21	3.09
7	2.56	22	3.14
8	2.64	23	3.14
9	2.71	24	3.14
10	2.76	25	3.15
11	2.77	26	3.19
12	2.81	27	3.44
13	2.82	28	3.46
14	2.89	29	3.54
15	2.93	30	3.56

Lower  $\bar{X}$  dispersion factors describe galaxies in which enemy units are closer together; thus, finding the enemy is more easily accomplished, and less mission time should be spent on search activities. Higher  $\bar{X}$  dispersions require that the captain make more strategic choices in conducting a search pattern. The enemy are spread thinly throughout the galaxy, and

these distributions generally result in a more difficult missions. Thus, use of the  $\bar{X}$  dispersion statistic provides a means to operationally define this variable in controlling task difficulty level.

2. Subjects: Six male and six female volunteers participated in this experiment. All subjects had been trained previously to Venture Captain Level 1 performance criteria, and each met U.S. Army enlistment standards. The purpose of this study and the procedures to be used were explained fully as required in the original Statement of Informed Consent signed by each subject on entry into the project. Subjects were paid \$3.50/hr for their participation.

3. Apparatus and procedures: The 12 subjects participated in three, 2-hr sessions. Sessions were conducted at the same time of day for a subject to control for circadian effects on performance. The apparatus and equipment configuration were identical to that described above in Section III-C-2 (Operating environment). In the first session, subjects performed missions using the training galaxies to assure that they had retained Level 1 performance values. Over the next two sessions, they performed missions in nine test galaxies varied for difficulty level. At the start of the second and third sessions, subjects performed a practice, or warm up, mission before performing the test missions. After each test mission subjects completed the SAM94 subjective fatigue and workload rating form, and no mission debriefing report was provided.

The nine test galaxies were constructed as shown below in Table 5, and presented in counterbalanced order.

TABLE 5

CHARACTERISTICS OF TEST GALAXIES USED IN DIFFICULTY LEVEL EXPERIMENT

	Star Distribution								
	Single			Double			Triple		
$\bar{X}$ Xenoid dispersion	Easy (2.06)	Med (2.70)	Hard (3.26)	Easy (1.96)	Med (2.86)	Hard (3.32)	Easy (1.95)	Med (2.77)	Hard (3.35)

4. Results: Each of the variables measured during performance was submitted to sex by star density by Xenoid dispersion ( $2 \times 3 \times 3$ ) analysis of variance (ANOVA) with repeated measures on the last two variables. The Greenhouse-Geisser correction was used to correct for inflated degrees of freedom due to correlations among repeated measures. Post hoc tests were conducted with Bonferoni  $t$ , which is especially suitable for repeated measures designs. Alpha was set at  $p < .05$  for all tests. No three-way interactions were observed.

The first question addressed was whether subjects' subjective ratings changed as a function of changes in star density and/or Xenoid dispersion. These data are summarized in Tables 6 and 7. It is apparent that star density significantly affected ratings of workload, stress, confidence and efficiency; as star density changed, these measures all changed in the expected direction. Changes in Xenoid dispersion were not related to changes in any subjective measures. No sex differences and no significant interactions were observed for these variables. The primary determinant of subjective change is therefore star density.

TABLE 6

CHANGES IN SUBJECTIVE RATINGS AS A FUNCTION OF DIFFICULTY LEVEL

Variable	Star Density		Xenoid Dispersion <sup>a</sup>	
	F	p <	F	p <
SAM94 Workload Rating	7.37	.05	2.53	ns
Postmission Workload:	5.32	.05	0.64	ns
Confidence	11.60	.01	0.04	ns
Efficiency	7.18	.05	0.14	ns
Stress	3.38	ns	0.23	ns
Fatigue	1.66	ns	0.22	ns
Pre- minus Postmission:				
Confidence	5.86	.05	0.27	ns
Efficiency	6.65	.05	0.17	ns
Stress	7.45	.05	1.56	ns
Fatigue	3.47	ns	0.49	ns
SAM94 Fatigue Rating	0.90	ns	0.14	ns

<sup>a</sup> No differences as a function of subject sex, and no significant interactions were found for these measures.

TABLE 7

## TABLE OF MEANS SIGNIFICANT F VALUES, SUBJECTIVE RATINGS

Variable	Star Density		
	Single	Double	Triple
SAM94 Workload Rating	3.17	3.50	3.75
Postmission:			
Workload	3.45	3.97	4.11
Confidence	7.08	6.40	6.18
Efficiency	6.75	6.20	5.92
Pre- minus Postmission:			
Confidence	-.50	.06	.19
Efficiency	-.39	-.03	.39
Stress	.11	-.14	-.39

The next question addressed was whether these subjective perceptions were also reflected in the major performance measures. Means for significant effects are shown in Table 8. Time per Xenoid increased significantly as star density increased ( $F = 54.18$ ,  $p < 0.01$ ) and as the Xenoid distribution became more dispersed ( $F = 12.46$ ,  $p < .01$ ). There was no significant interaction between the two. Energy per Xenoid increased with star density ( $F = 17.90$ ,  $p < .01$ ) and with Xenoid dispersion ( $F = 5.11$ ,  $p < .05$ ). Figure 7 shows the effects of difficulty level factors on time and energy per Xenoid. Overall time per command showed a similar pattern (Star density  $F = 17.18$ ,  $p < .01$ ; Xenoid dispersion  $F = 3.76$ ,  $p < .10$ ). No sex differences and no significant interactions were observed for these variables. Again, star density appears to be the most important variable in determining mission difficulty.

Variables reflecting the internal structure of mission performance were then examined to determine how these differences came about. Theoretically, star density should have a direct effect on navigation, requiring the captain to maneuver more and to use more fuel to complete the mission. Similarly, Xenoid dispersion should increase search behavior and require the captain to make more decisions in conducting an effective search strategy. The effects of task difficulty on variables related to navigation and search strategy are shown in Tables 9 and 10; data related to search strategy are presented in Tables 11 and 12. Figure 8 presents illustrative data on two of the major variables. Altering star density clearly affected the captain's navigation strategy, particularly with regard to percent of all commands used for navigation, the number of navigation commands, and the fuel units and commands used for maneuvering. Xenoid dispersion, on the other hand, had its greatest effects on navigation items relevant to searching out the enemy.

TABLE 8

TABLE OF MEANS SIGNIFICANT CHANGES IN MAJOR PERFORMANCE VARIABLES

<u>Star Density</u>			
<u>Variable</u>	<u>Single</u>	<u>Double</u>	<u>Triple</u>
Time per Xenoid (sec)	31.06	44.09	49.39
Energy per Xenoid (units)	181.2	217.7	224.1
Overall Time/Command (msec)	7,174	8,342	8,632

<u>Xenoid Dispersion</u>			
	<u>Easy</u>	<u>Medium</u>	<u>Difficult</u>
Time per Xenoid (sec)	36.75	38.93	46.95
Energy per Xenoid	193.8	200.0	229.2

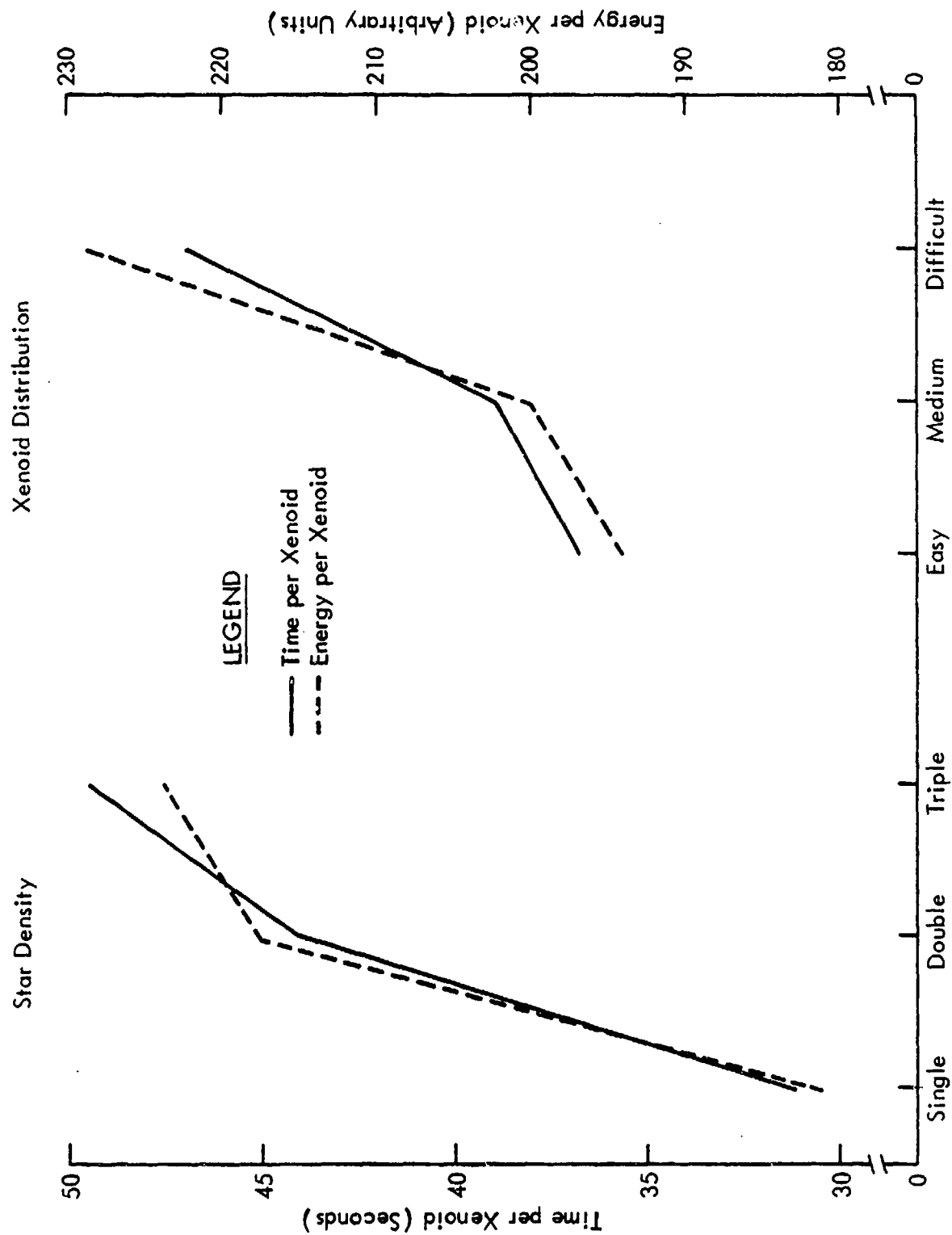


Figure 7 - Effects of Difficulty Level Factors on Time (Left Axis) and Energy (Right Axis) Utilization

TABLE 9

EFFECTS OF TASK DIFFICULTY ON VARIABLES RELATED TO NAVIGATION

<u>Variable</u>	<u>Star Density</u>		<u>Xenoid Dispersion</u>	
	<u>F</u>	<u>p &lt;</u>	<u>F</u>	<u>p &lt;</u>
Percent of all commands that were used for navigation	67.32	.001	2.11	ns
Number of navigation commands	75.07	.001	45.20	.001
Percent of all navigation commands used to:				
Search	1.90	ns	95.41	.001
Attack	77.30	.001	68.56	.001
Maneuver	118.60	.001	5.19	.05
Number of navigation commands for:				
Search	12.76	.01	111.46	.001
Attack	0.70	ns	8.34	.05
Maneuver	106.27	.001	4.28	ns
Evade	1.42	ns	2.51	ns
Fuel used in navigation to:				
Search	9.56	.05	156.18	.001
Attack	5.90	.05	3.54	ns
Maneuver	45.38	.001	3.50	ns
Evade	0.88	ns	1.89	ns

TABLE 10

TABLE OF MEANS  
SIGNIFICANT VARIABLES RELATED TO NAVIGATION

<u>Star Density</u>			
<u>Variable</u>	<u>Single</u>	<u>Double</u>	<u>Triple</u>
% of all commands used for navigation	27.54	30.97	34.93
Number of navigation commands	23.42	32.97	39.39
% of all navigation commands used for:			
Attack	54.68	39.54	31.78
Maneuver	24.47	37.88	47.34
Number of navigation commands for:			
Search	4.75	6.95	6.80
Maneuver	5.67	12.56	18.75
Fuel used in navigation for:			
Search	398.9	555.6	572.8
Attack	627.2	615.7	596.1
Maneuver	44.9	118.2	169.0

<u>Xenoid Dispersion</u>			
	<u>Easy</u>	<u>Medium</u>	<u>Hard</u>
Number of navigation commands	25.69	32.56	37.53
% of all navigation commands used for:			
Search	6.20	16.10	30.15
Attack	51.46	43.10	31.45
Maneuver	39.45	36.60	33.63
Number of navigation commands for:			
Search	1.85	5.57	11.08
Attack	12.32	12.28	11.42
Fuel used in navigation for:			
Search	139.7	453.9	933.6

TABLE 11

## EFFECTS OF TASK DIFFICULTY ON VARIABLES RELATED TO SEARCH STRATEGY

Variable	Star Density		Xenoid Dispersion	
	<u>F</u>	<u>p &lt;</u>	<u>F</u>	<u>p &lt;</u>
Percent of total commands used for:				
Short-range scans (SRS)	6.41	.05	1.26	ns
Long-range scans (LRS)	24.20	.001	0.38	ns
Long-range scan histories (LRSH)	0.77	ns	24.44	.001
LRS + LRSH	6.84	.05	37.31	.001
Number of:				
SRS	18.33	.01	61.59	.001
LRS	12.96	.01	43.28	.001
LRSH	5.42	.05	41.06	.001
Number of LRS with no new information	2.78	ns	21.69	.001
Number of LRSH with no new information	1.62	ns	4.98	.05
Time between LRSH commands	2.10	ns	6.44	.05
Number of LRSH/No. of LRS x 100	2.56	ns	4.73	.05
Percent resource commands, fuel	0.74	ns	20.81	.005
Energy units allocated to fuel	2.88	ns	52.74	.001
Percent energy allocated, fuel	0.35	ns	29.04	.001
Percent quadrants of new information per LRS	2.17	ns	10.93	.01
$\bar{X}$ time to extract information per LRS	0.83	ns	12.50	.01

TABLE 12

TABLE OF MEANS  
SIGNIFICANT VARIABLES RELEVANT TO SEARCH STRATEGY

<u>Star Density</u>			
<u>Variable</u>	<u>Single</u>	<u>Double</u>	<u>Triple</u>
% of total commands used for:			
Short-range scans (SRS)	19.2	18.7	17.9
Long-range scans (LRS)	19.0	18.6	16.5
LRS + LRSH	26.7	27.0	25.0
Number of:			
SRS	16.5	19.9	20.1
LRS	16.2	19.7	18.6
LRSH	8.4	11.0	11.2
<u>Xenoid Dispersion</u>			
	<u>Easy</u>	<u>Medium</u>	<u>Hard</u>
% of total commands used for LRSH	7.3	8.8	11.3
% of total commands used for LRS + LRSH	24.0	26.0	28.8
Number of:			
SRS	15.2	18.9	22.5
LRS	14.6	18.1	21.8
LRSH	6.7	10.0	13.9
Number of LRS with no new information	2.1	3.2	4.9
Number of LRSH with no new information	1.5	2.3	3.4
Number of LRSH/No. of LRS x 100	48.1	54.8	65.0
Percent resource commands, fuel	26.5	40.8	59.6
Energy units allocated to fuel	453.6	657.4	1,060.4
Percent energy allocated, fuel	23.2	32.2	43.0
Percent quadrants of new information/LRS	86.5	83.9	78.8
$\bar{X}$ time to extract information/LRS	4.5	3.9	3.6
Time between LRSH commands	95.9	89.5	71.1

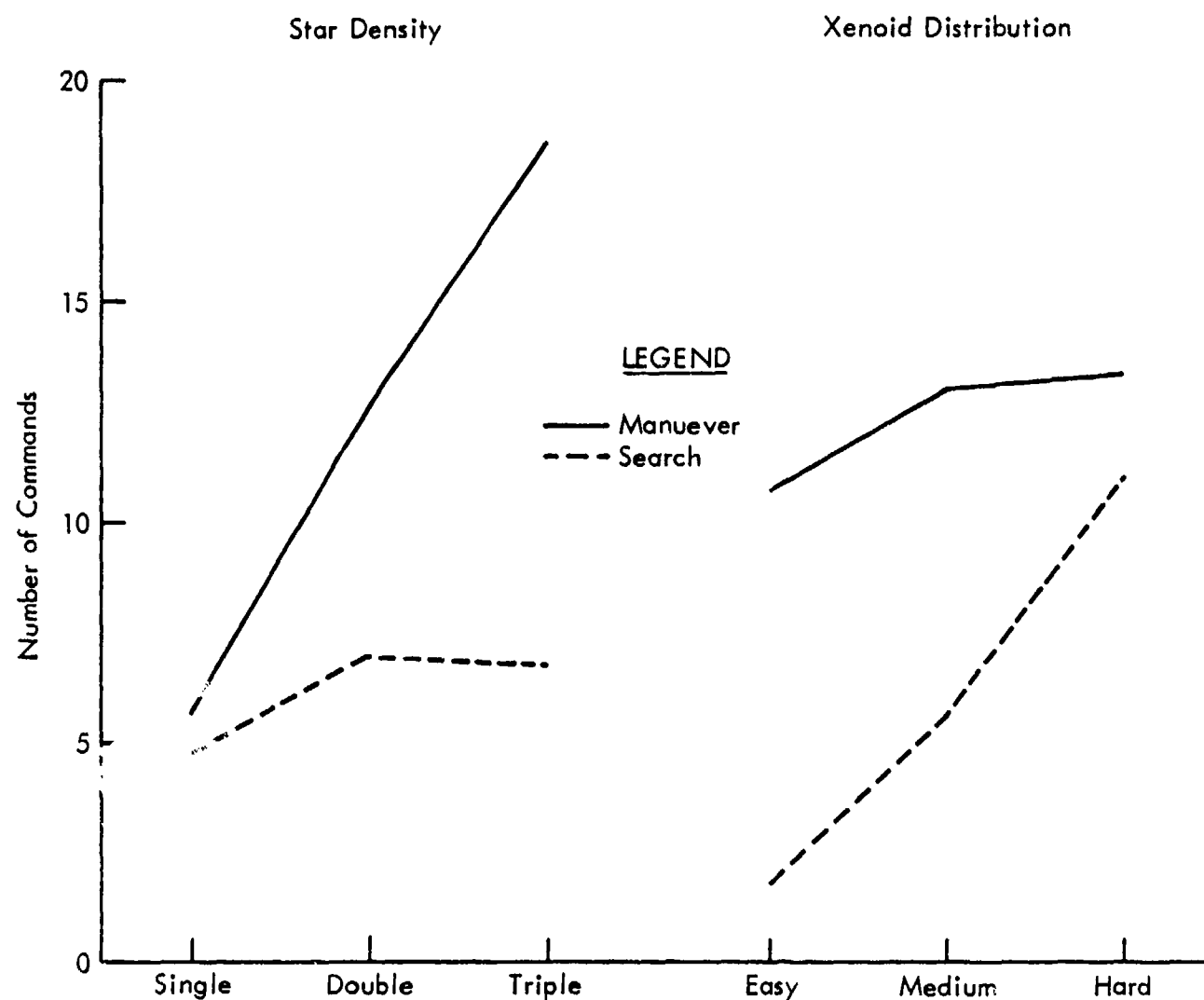


Figure 8 - Effects of Star Density and Xenoid Distribution  
on Maneuver and Search Commands

In terms of search strategy, both star density and Xenoid dispersion increased the amount of scanning behavior, as measured by the number of scans performed, although this effect was much greater for Xenoid dispersion. Star density also affected percent of commands used for scanning, but this result is undoubtedly due to the general increase in number of commands associated with increased star density. All other search strategy variables were affected only by Xenoid dispersion. The results shown in Tables 11 and 12 clearly reflect the more complex search strategy required to seek out Xenoids which are widely scattered in the galaxy. It is particularly interesting to note that the efficiency of information seeking declines as Xenoid dispersion increases. The number of scans containing no new information increases with increasing Xenoid dispersion, while the mean time to extract information from the scans and the time between scans decreases.

Tables 13 and 14 present the results obtained when effects of task difficulty on offensive strategy were examined. As star density increased, the mean quadrant kill time (time from entering the quadrant to destroy all enemy, computed over all quadrant types) increases. Analysis of other offensive strategy measures helps explain this decline in performance. The number of moves to a better battle position increased, probably reflecting the tactical advantages obtained by "hiding" behind stars while attacking the enemy. Star density also affected subjects' choice of weapon systems; torpedo commands decreased with increased star density, while phaser commands increased. This finding is consistent with the tactical advantages often obtained by using phasers when some enemy are "behind" stars. This interpretation is supported by the fact that star density had no effect on correctness of weapon choice. Finally, as discussed below, the accuracy with which subjects used either weapon system declined as star density increased.

The factor of Xenoid dispersion was not expected to have an effect on offensive strategy and this was confirmed. Dispersion is important only when attempting to find the enemy; once the captain enters a quadrant containing enemy cruisers, how the enemy is dispersed over other quadrants is not relevant. Tables 13 and 14 show a number of significant findings relevant to Xenoid dispersion and offensive strategy. These are all expressed in terms of percent changes. Since the total number of commands increases with increases in dispersion ( $F = 66.27$ ,  $p < .001$ ), those results are artifacts arising from a change in a different mission parameter. No differences in the absolute number of offensive commands or units of fuel used offensively were found to be related to Xenoid dispersion.

Figure 9 shows the effects of star density and Xenoid dispersion on three major defensive strategy variables. As Xenoid dispersion increased, the mean level of shield energy on entering enemy quadrants increased linearly ( $F = 12.09$ ,  $p < .005$ ). Star density affected both mean shield level during weapon commands and the percent of weapon commands issued with inadequate shields. As star density increased, subjects became more efficient in using the defensive shields to protect their cruisers.

TABLE 13

EFFECTS OF DIFFICULTY ON VARIABLES RELATED TO OFFENSIVE STRATEGY

<u>Variable</u>	<u>Star Density</u>		<u>Xenoid Dispersion</u>	
	<u>F</u>	<u>p &lt;</u>	<u>F</u>	<u>p &lt;</u>
$\bar{X}$ quadrant kill time	7.50	.05	0.95	ns
Number of moves to better battle position	10.33	.01	.72	ns
Percent correct weapon choice	1.68	ns	0.32	ns
Phaser use:				
Number of phaser commands	10.13	.01	0.36	ns
Percent commands, phasers	10.58	.01	44.66	.001
Percent resource commands, phasers	0.18	ns	10.15	.01
Energy units allocated to phasers	4.56	ns	0.63	ns
Percent energy allocated, phasers	0.77	ns	23.46	.001
Torpedo use:				
Number of torpedo commands	5.01	.05	1.69	ns
Percent commands, torpedos	7.99	.05	13.53	.005
Percent resource commands, torpedos	1.91	ns	3.93	ns
Energy units allocated to torpedos	1.19	ns	2.04	ns
Percent energy allocated, torpedos	1.05	ns	3.27	ns
Percent torpedoes on target	4.60	.05	2.38	ns
Percent total commands used for weapons	22.65	.001	74.60	.001
Percent commands for offensive resupply	3.32	ns	13.30	.01
Percent of total commands used for attack	37.58	.001	78.54	.001

TABLE 14  
TABLE OF MEANS FOR SIGNIFICANT VARIABLES RELATED  
TO OFFENSIVE STRATEGY

<u>Star Density</u>			
	<u>Single</u>	<u>Double</u>	<u>Triple</u>
Mean quadrant kill time	14.5	20.2	23.7
Number of moves to better battle position	.3	.9	2.0
Percent of total commands used for attack	14.3	11.7	10.7
Phaser use:			
Number of phaser commands	11.2	12.0	13.4
Percent commands, phasers	14.0	11.8	12.3
Torpedo use:			
Number of torpedo commands	2.4	2.4	1.4
Percent commands, torpedos	2.9	2.4	1.3
Percent torpedoes on target	96.9	86.9	85.0
Percent total commands used for weapons	16.2	13.7	13.1
<u>Xenoid Dispersion</u>			
	<u>Easy</u>	<u>Medium</u>	<u>Hard</u>
Percent of total commands used for attack	14.6	12.6	9.5
Phaser use:			
Percent commands, phasers	14.7	12.9	10.4
Percent resource commands, phasers	57.4	47.0	41.8
Percent energy allocated, phasers	66.1	61.6	52.4
Torpedo use:			
Percent commands, torpedos	3.0	2.0	1.5
Percent total commands used for weapons	17.0	14.4	11.6
Percent commands for offensive resupply	4.8	3.5	3.4

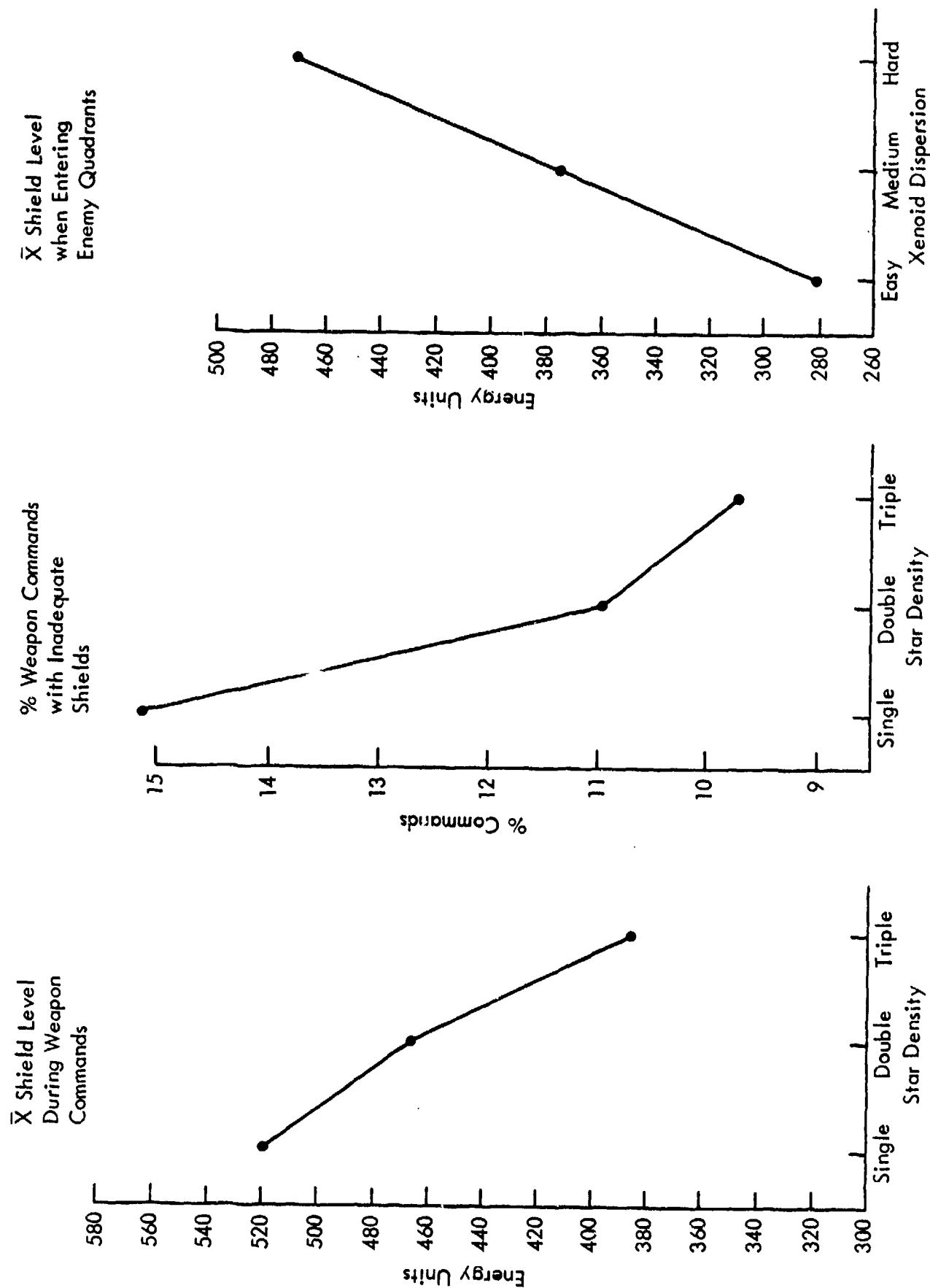


Figure 9 - Effects of Star Density on Xenoid Dispersion on Defensive Strategy

Shield levels during weapon commands decreased ( $F = 12.55$ ,  $p < .005$ ), but this decrease had little operational significance, since levels were typically maintained above the minimum needed to protect the Venture, and the percent of weapon commands issued with inadequate shield protection tended to decline with increasing density ( $F = 3.58$ ,  $p < .10$ ).

The correct calculation of torpedo courses, navigation courses, and phaser payloads are primary measures of information processing. While star density had no effect on navigation course calculation accuracy, there was a significant decline in the accuracy of torpedo course calculations. The percent of torpedoes on target decreased as star density increased ( $F = 4.60$ ,  $p < .05$ ). Phaser payload calculation accuracy also decreased as star density increased ( $F = 9.73$ ,  $p < .05$ ), and the percent of successful phaser attacks decreased ( $F = 6.76$ ,  $p < .05$ ). Increasing star density appears to have a deleterious effect on information processing operations during the tactical segments of a mission. Xenoid dispersion, as expected, had no effect on variables associated with this situation.

Neither recent nor long-term memory as measured during docking procedures was affected by star density or Xenoid dispersion. However, errors of memory storage during command operations (i.e., the number of attack and navigation commands without adequate resources to carry out the command) increased as Xenoid dispersion increased ( $F = 5.61$ ,  $p < .01$ ). This distinction between docking and mission performance is also seen in psychomotor skills. Phantom response time became slower as Xenoid dispersion increased ( $F = 6.38$ ,  $p < .05$ ); tracking performance was not affected by difficulty level.

The appropriate and efficient allocation of energy from reserves to operational systems, and the minimization of dockings at the HUB to replenish resources are important elements of the task. The number of commands used to reallocate resources, and the total energy allocated, were affected by both star density and Xenoid dispersion. The number of resource commands increased as star density increased ( $F = 6.97$ ,  $p < .05$ ), and as Xenoid dispersion increased ( $F = 13.01$ ,  $p < .01$ ). Amount of energy allocated was greater for double and triple star densities than for single densities ( $F = 6.86$ ,  $p < .05$ ) and increased as Xenoid dispersion increased ( $F = 10.81$ ,  $p < .01$ ). The number of dockings at HUB was not affected by Xenoid dispersion, but did increase as star density increased ( $F = 7.24$ ,  $p < .025$ ).

The experiment was designed to allow direct comparison of the performance of male and female subjects. Table 15 shows the variables on which significant main effects for sex of subject were obtained, and Table 16 presents significant interactions between sex and Xenoid dispersion. No significant interactions between sex and star density occurred. It is interesting to note that no sex differences were found on major performance variables. Female subjects made more navigation and maneuver commands, moved more frequently to better battle positions, and used more

TABLE 15

VARIABLES SHOWING SIGNIFICANT MAIN EFFECTS FOR SEX OF SUBJECT

<u>Variable</u>	<u>F</u>	<u>p &lt;</u>	<u>Means</u>	
			<u>Male</u>	<u>Female</u>
Number of moves to better battle position	7.63	.05	.5	1.6
Fuel used for attack	10.34	.01	575.4	632.6
Number of navigations for maneuver	7.90	.05	11.07	13.57
Number of navigation commands	8.99	.05	29.46	34.39
Number of short-range scans	14.97	.01	16.78	20.91
Number of long-range scans	9.33	.05	16.44	19.93
Percent of navigations for attack	11.32	.01	45.09	38.91
Percent of resource commands, life support system	6.50	.05	3.99	12.29
Number of informational requests which produced no new information	5.38	.05	3.61	7.89
Number of voluntary dockings	7.54	.05	0.87	1.13
Number of life support resupplies	7.41	.05	0.28	1.17
Recent memory, pre minus post response time	7.79	.05	11.5	8.5
Number of commands per mission	9.06	.05	94.7	115.6
Percent commands for weapons	8.40	.05	15.4	13.3
Percent commands for defensive resupply	8.42	.05	.4	1.2

TABLE 16

SIGNIFICANT INTERACTIONS BETWEEN SEX AND XENOID DISPERSION

	<u>F</u>	<u>p</u>	<u>Means</u>					
			<u>Males</u>			<u>Females</u>		
			<u>Easy</u>	<u>Medium</u>	<u>Hard</u>	<u>Easy</u>	<u>Medium</u>	<u>Hard</u>
% long-term memory correct	6.62	.05	47.2	52.8	16.7	36.1	44.4	50.0
No. of LRS, no new information	5.10	.05	1.6	2.5	3.2	2.6	3.8	6.7
% commands for phasers	5.84	.05	15.0	14.6	11.0	14.4	11.1	9.9
% resource commands, phasers	6.25	.05	56.4	57.0	41.2	58.3	37.0	42.4
% resource energy, phasers	7.71	.05	66.2	68.9	52.9	66.1	54.4	51.9
% navigation commands, attack	5.93	.05	15.5	13.6	10.3	13.8	10.6	8.7

fuel for attacks. They made greater use of short- and long-range information sources than males. Males used a greater proportion of their navigation commands for attack purposes, made fewer reallocations of energy to their life support systems, and docked at the HUB for resupply less frequently.

The significant interactions presented in Table 16 suggest that females were less affected by Xenoid dispersion than males in terms of percent memory correct, but more affected on measures of phaser usage and number of long-range scans containing no new information. Figure 10 illustrates these results.

5. Discussion: The goal of this initial study of STAR was to establish three levels of task difficulty. This goal was accomplished. Variation in the selected task parameters of star density and Xenoid dispersion resulted in the hypothesized effects on subjective measures and mission performance variables. Moreover, the findings obtained were robust, internally consistent, and readily interpretable. These characteristics of the data are highly encouraging, particularly in light of the complexity of the task performed, the relatively small number of subjects tested, and the large number of variables assessed.

The primary factor underlying task difficulty is star density. Increases in density resulted in increases in ratings of workload and stress, and decreases in reported levels of confidence and efficiency. These changes in subjective report were paralleled by corresponding changes in mission performance measures. With increasing density, subjects required more time to destroy the enemy, and used more energy to accomplish their mission goal. Analysis of specific command structures and sequences demonstrated, as predicted, that the density factor had its greatest impact in the area of navigation. This phenomenon was reflected not only in the number of navigation commands issued, but also in the specific use of these commands and in the amount of fuel expended.

Star density also had a major impact on tactical decision making. The subjects were quick to take advantage of the "cover" afforded by increasing star density, as reflected by an increase in the number of navigations to a better battle position during attack. They also reevaluated the relative merits of the two weapon systems available to them. Torpedo use decreased, and phaser use increased. Given their situation, this was a correct decision, and was reflected by the finding that increasing density did not result in a decrease in correct weapon choice. However, increasing density did have a significant effect on the information processing time required to make tactical decisions, and the types of errors committed. The time required to destroy the enemy over all types of enemy configurations encountered increased directly with density. The percent of torpedoes on target, the number of correct phaser payload calculations, and the percent of successful phaser attacks showed similar declines with increasing density. It is interesting to note that calculating a torpedo course requires exactly the same operation as calculating a navigation course, yet navigation calculations were not affected by increasing density. Thus, although subjects adjusted correctly to the changes in density, these

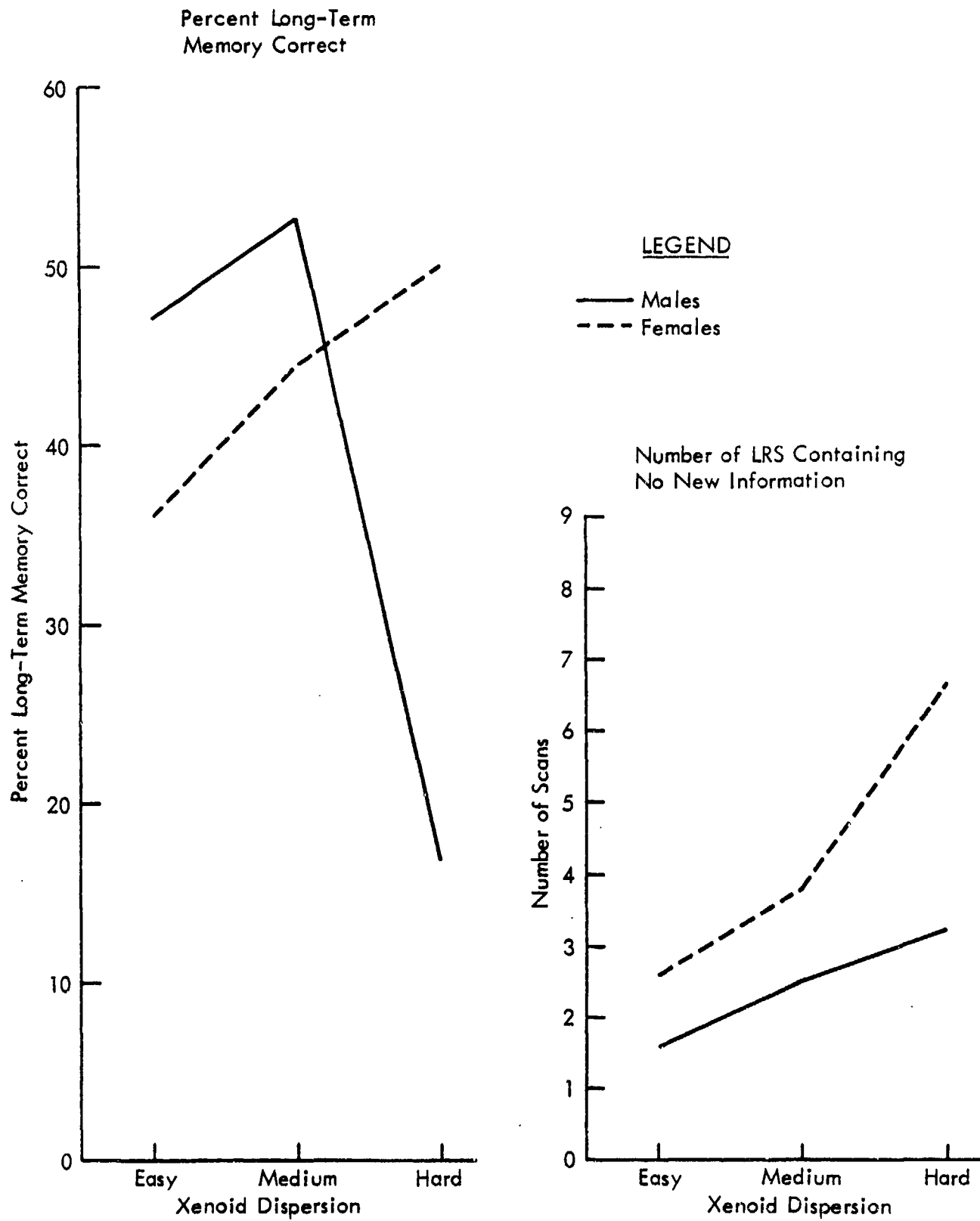


Figure 10 - Interactions Between Sex and Xenoid Dispersion

adjustments took time and resulted in significantly increased tactical error rates.

The factor of Xenoid dispersion was expected to have very different effects compared to star density. In one sense, density corresponds to clutter. Increases in density are very obvious to the subjects. They do not know where the enemy is to begin with, so increased dispersion is simply not apparent. As the mission progresses, however, increased dispersion represents more of a challenge than a frustration. This may explain why changes in dispersion were not related to changes in any subjective measure.

Dispersion did not have a major impact on information processing aspects of mission performance. As expected, alterations in Xenoid dispersion had their greatest effect on measures related to search decision strategy. With increasing dispersion, subjects significantly increased the use of all available information systems. The time spent in searching for the enemy and fuel used for search markedly increased.

Analysis of the fine structure of information-seeking activity can help determine the relative efficiency of the strategies employed, and reveal the common types of errors that occur as dispersion increases. Although subjects requested more information, they spent significantly less time extracting relevant information from the displays they requested. The number of times their requests for information resulted in no new information at all increased markedly as dispersion increased. The same phenomenon is seen in the kinds of information subjects requested. Subjects make much greater use of the LRSH, which is designed to be a major decision aid in search activity, and the proportion of LRSH to LRS increased markedly. Thus, the pattern observed with increasing Xenoid dispersion is a marked increase in information seeking activity; this activity can be characterized as being more redundant, less thorough in extracting relevant information, and less efficient.

Use of these two game parameters to control task difficulty, therefore, can provide important information about specific ways in which changes in task difficulty affect the performance of a mission. This is clearly demonstrated by examination of the ways in which men and women go about performing STAR. Overall, no sex differences were found on major performance variables, and men and women did not respond differently to increases in star density. Men and women, however, did appear to perform their missions differently. Women made greater use of information sources, attended more to resupplying their ships, issued more commands, and performed more navigation than men. Most of the interactions between sex and Xenoid dispersion can be explained by these stylistic differences. Thus, on the basis of this small unmatched sample of six men and six women, there appears to be a difference in the process of performing this complex task, but no difference in the end product of the task. STAR might prove to be a valuable tool for use in future studies of individual differences in complex human performance, and of the kinds of errors that might be expected to occur with increasing difficulty or stress.

### C. Evaluation of Specific Stressors on STAR Performance

1. Introduction: The future utility of STAR would be greatly enhanced if it were found that variations in selected task parameters could be used to produce a crisis, or very high stress, condition for the subject. The task could then be used to evaluate the impact of sustained performance demands, environmental stressors, etc., on cognitive functioning under crisis and noncrisis conditions. Consequently, the purpose of this study was to first identify likely parameters for inducing crisis, and then to evaluate performance under controlled conditions that systematically varied stress and task difficulty parameters.

We had originally proposed to induce performance stress by placing a time limit on decision making. In essence, the captain would be required to make a decision every so many seconds during the mission. If no decision was made within the specified time period, the program would penalize the captain by draining energy from the reserve supply of the VENTURE at a fixed rate per unit time until a decision was made and a command entered at the terminal. Our experience with STAR suggested that this was not an optimal approach. Limiting decision time would distort the time dependency characteristics of a number of STAR measures. In addition, this type of approach to inducing stress is basically nondiscriminative and easily circumvented by intelligent subjects. Our analysis of the training data and of performance in the previous difficulty level study suggested that two major aspects of the mission could be used to induce stress. The first was overall mission time, and the second was the amount of resources the captain had available to accomplish the mission.

At the start of training, subjects experienced a good deal of stress because they could not make correct decisions fast enough to destroy all the Xenoids present in the time allowed for the mission. They cursed and yelled and groaned, made mistakes, complained about "computer errors," and wanted to know how their performance compared to that of others. In short, they showed many of the outward signs of being in a performance crisis that was personally relevant to them. With practice, however, their skill increased. When they reached Level 1 performance criteria, they could easily destroy all Xenoid cruisers in the first 15 min of a 30-min mission. The behavioral signs and subjective ratings of performance stress decreased with skill acquisition. These signs reappeared only when the 15-min mark in a mission (the time criterion for Level 1 performance) approached, and they still had several enemy cruisers yet to find and destroy. Thus, it seemed likely that alterations in mission time could be used to induce stress without at the same time interfering with the time dependency characteristics of specific STAR measurement parameters.

The availability of resources seemed to be another likely stressor. During training, subjects basically had unlimited resources available to accomplish a mission. In the initial training missions they had 3,000 units of energy available in their reserve supply on board VENTURE, and could allocate the energy to various systems as they wished. When this supply ran out, they could dock at the HUB and obtain a complete resupply of energy

and armament. They had to learn to meet the energy use criterion for Level 1 performance, but no restrictions were placed on the total supply they had to draw from.

After analysis of the initial training data, we decided for a variety of reasons to change the energy allotments available on board VENTURE. Instead of 3,000 reserve units, the captain now had only 1,500 units. Corresponding changes were also made in other system allotments. Unlimited docking at the HUB was still allowed during a mission. All subjects trained under these new restrictions. Subjects found that they had to pay much closer attention to how and when they used their resources. They made more performance errors (e.g., attempting to navigate or use weapons without adequate energy available), had their cruisers blown up more frequently, dropped in rank, and the outward signs of stress reappeared. Eventually, new skills were acquired, Level 1 was reattained, and stress signs decreased in frequency and intensity.

A highly interesting and potentially useful aspect of these observations is that stress under these circumstances seemed to be due primarily to the inability of subjects to adjust to and function differently in a new situation. If they had simply decided to dock at the HUB more frequently, they would have had more than enough energy to conduct a mission. In fact, the primary reason we changed the energy allocation scheme was to force more dockings so that we could collect additional data during a mission. Subjects did, in fact, dock more often, but only after going to inordinate lengths attempting to function with their smaller onboard allotments. This observation is useful because it suggests that stress can be induced by only apparent reductions in the total resources available. In addition, the performance measures affected by resource reduction seemed to be different from those affected by reduction in overall mission time. Thus, the use of these two parameters, either alone or in combination, might prove valuable in producing a stress situation with wide ranging impact on performance. When employed in conjunction with variations in task difficulty level, the possibility of producing a crisis seemed feasible.

2. Subjects: The six male and six female volunteers who had participated in the previous difficulty level study served as subjects in the present study. Each was informed that "various changes had been made in STAR that might make the game more interesting and challenging," and that we wanted to evaluate the impact of these changes on performance. Subjects were paid \$3.50 for each hour of participation.

3. Experimental design and procedures: Twelve subjects performed two practice missions and eight test missions over two 2-hr sessions using the equipment configuration described previously. Sessions were conducted at the same time of day for a subject, and each session began with a practice mission. The eight test missions were presented in counterbalanced order across subjects.

Two difficulty levels were employed to examine interactions between crisis conditions and task difficulty. The easy difficulty level was defined as a galaxy with standard star density and a X Xenoid dispersion

factor between 1.95 and 2.05. The hard difficulty level was defined as a galaxy with triple standard star density and a  $\bar{X}$  dispersion factor between 3.25 and 3.35. The two control galaxies were duplicates of the galaxies used in the previous difficulty level study. Performance in these two galaxies provided a control condition for the present study, and also a means to evaluate the test-retest reliability of all STAR measures over different testing sessions. Table 17 shows the experimental design of the present study, and Table 18 presents the major characteristics of the eight test mission/galaxy configurations.

Data from the previous difficulty level study were evaluated to establish the time and energy values shown in Table 18. The mean and standard deviation of time and energy values obtained over 48 previous missions were computed. For this study, time and energy restrictions in the experimental missions were set at values equal to the mean plus one standard deviation. In essence, if these limits had been imposed in the previous study, it would have meant that about 12% of the previous missions would have been unsuccessful (6 of 48 missions).

Use of the energy restriction made it necessary to add one additional command to STAR. This was a surrender command. This was necessary because if subjects ran out of energy by not allocating their onboard energy correctly, and could not dock to obtain additional supplies, their only option would be to surrender. The operation of this command was explained to subjects at the start of the study.

4. Results: The two galaxies used for control missions in this study were duplicates of an easy and a hard difficulty level galaxy used in the previous experiment. Thus, the first analysis compared performance on these galaxies over the two experiments in order to determine if the instructions used in the stress study had any appreciable impact on performance. Student's  $t$  was performed separately for each of the 92 measure obtained during a mission. For the easy galaxies, only two measures showed significant differences. Subjects showed a greater increase in stress ( $t = 2.80$ ,  $p < .02$ ) and took less time between LRSH ( $t = 3.24$ ,  $p < .01$ ) during the stress study. For the hard galaxies, only one of the 92 comparisons was significant. Subjects issued a higher percent of their commands for LRS during the stress study ( $t = 2.27$ ,  $p < .05$ ). Given that 184  $t$  tests were performed, and only three produced significant differences, performance on these control galaxies during the two experiments was essentially the same.

If the experimental paradigm was successful in producing a crisis situation, we would expect higher subjective ratings of stress and workload for missions with time and energy limitations than for control missions. Results are shown in Table 19. As predicted, the combined time and energy restricted missions resulted in the largest increases in ratings of workload and stress, confirming the success of the paradigm.

TABLE 17

EXPERIMENTAL DESIGN OF THE STRESS STUDYEasy Difficulty Level Missions

Control	Time (T) Restriction	Energy (E) Restriction	T + E Restriction
---------	-------------------------	---------------------------	----------------------

Hard Difficulty Level Missions

Control	Time (T) Restriction	Energy (E) Restriction	T + E Restriction
---------	-------------------------	---------------------------	----------------------

TABLE 18

CHARACTERISTICS OF EIGHT TEST MISSIONS USED IN THE STRESS STUDY

<u>Galaxy Number</u>	<u>Difficulty Level/ Restriction Type</u>	<u>Mission Time (min)</u>	<u>Reserve<sup>b</sup> Energy (units)</u>	<u>Number of<sup>c</sup> Dockings Allowed</u>
<u>Easy</u>				
811	Control <sup>a</sup>	30	1,500	Unlimited
812	Energy	30	1,800	None
821	Time	12	1,500	Unlimited
822	Both	12	1,800	None
<u>Hard</u>				
911	Control <sup>a</sup>	30	1,500	Unlimited
912	Energy	30	3,400	None
921	Time	22	1,500	Unlimited
922	Both	22	3,400	None

a This galaxy is an exact duplicate of a galaxy used in the previous difficulty level experiment.

b Reserve energy is the actual amount the subject has available on board VENTURE to allocate to all systems as the need arises. During energy restriction missions, the total amount was given at the start of a mission and the subject could not dock at the HUB to obtain more.

c Unlimited docking means that the subject is free to dock anytime during a mission. The energy levels in all VENTURE systems, including reserve energy, are completely restored to their original levels at each docking.

TABLE 19

EFFECTS OF TIME AND ENERGY RESTRICTIONS ON SUBJECTIVE RATINGS

<u>Variable</u>	<u>F</u>	<u>p &lt;</u>	<u>Mean Rating</u>			
			<u>Control</u>	<u>Time</u>	<u>Energy</u>	<u>Both</u>
Workload	6.01	.05	3.4	3.7	3.7	4.4
Change in Stress <sup>a</sup>	5.86	.05	-.17	-.04	-.25	-.58
Change in Confidence <sup>a</sup>		NS				
Change in Efficiency <sup>a</sup>		NS				
Change in Fatigue <sup>a</sup>		NS				
SAM 94 Workload	4.70	.10	3.0	3.7	3.5	4.0
SAM 94 Fatigue		NS				

a Change scores were pre-mission ratings minus post-mission ratings; the higher the negative number, the greater the increase in that variable.

The effects of time and energy restrictions on the major parameters of the task were examined next. In the difficulty experiment, all Xenoids were destroyed in all missions. In the present experiment, subjects failed to destroy 32 enemy cruisers; 1 in time restricted missions, 2 in energy restricted missions, and 29 in missions with limitations on both time and energy. Similarly, subjects failed to protect their own cruisers on a significant number of missions. Thirty-four of the total of 96 VENTURE cruisers sent on missions were lost. Sixteen of these were lost in missions with combined time and energy restrictions.

Difficulty level by time restriction by energy restriction ( $2 \times 2 \times 2$ ) ANOVAs for repeated measures were performed. Surprisingly, no main effects of stress level were found for overall performance variables such as time per Xenoid, energy per Xenoid and overall time per command. Planned comparisons between control and combined stress galaxies revealed a further surprising phenomenon, and helped explain the lack of overall main effects. Contrary to expectation, the major impact of combined stress was on the easy galaxies rather than the hard galaxies. In the easy galaxies, the hypothetical effects of stress were apparent. Significantly more energy per Xenoid ( $t = 1.94, p < .05$ ) and time per Xenoid ( $t = 2.00, p < .05$ ) were expended on easy difficulty, combined stress galaxies. No such difference was found for the hard galaxies. Further examination of the number of unsuccessful missions and the number of Ventures destroyed revealed the same pattern.

As shown in Figure 11, this pattern is due at least in part to an increase in risk taking behavior under stress. For example, during the combined stress conditions in hard galaxies, 9% of weapon commands were issued without adequate shield protection. In contrast, 30% of the weapon commands in the easy, combined stress galaxies had inadequate shield protection. Similarly, Figure 12 shows the mean shield level on entering enemy quadrants. This figure demonstrates that during easy-level missions subjects did not maintain their shields at a level adequate to protect them from attack by even one enemy cruiser. When time stress was added to the mission, subjects continued to place their cruiser at risk. The situation was somewhat different in high difficulty level missions. Subjects maintained shield levels at a high, adequate level. With time stress, they let this level drop, but still maintained it at a level adequate to protect against an enemy attack. Energy restriction did not affect this particular risk-taking variable.

Figure 13 examines risk-taking behavior during attack situations, when the Venture is in greatest danger. When no time restrictions are in effect, there is a slight increase in shield levels during weapon commands; with time restriction, there is a slight decrease. This phenomenon is more dramatic when the difference between actual shield level and the minimum required to protect the Venture is examined. When both time and energy are restricted, the difference decreases almost 50% from control.

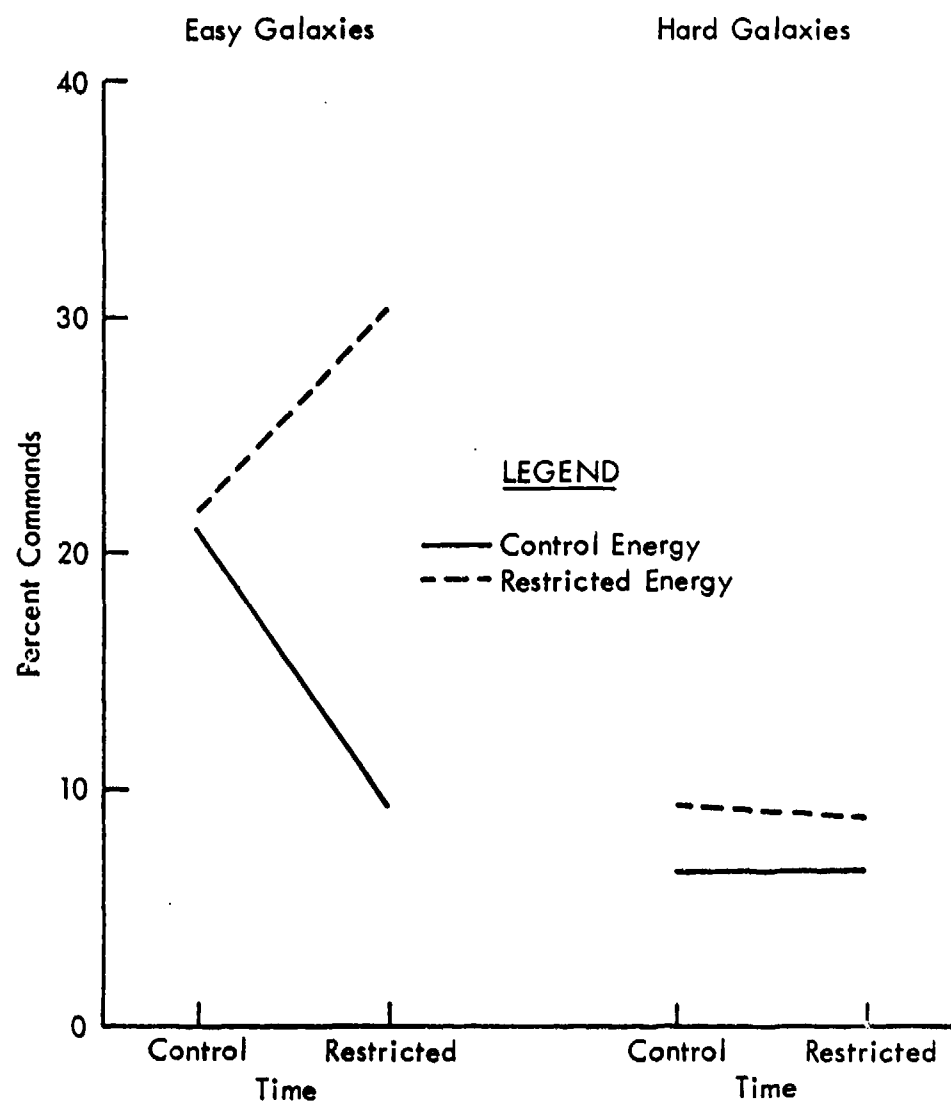


Figure 11 - Percent Weapon Commands Issued Without Adequate Shield Protection

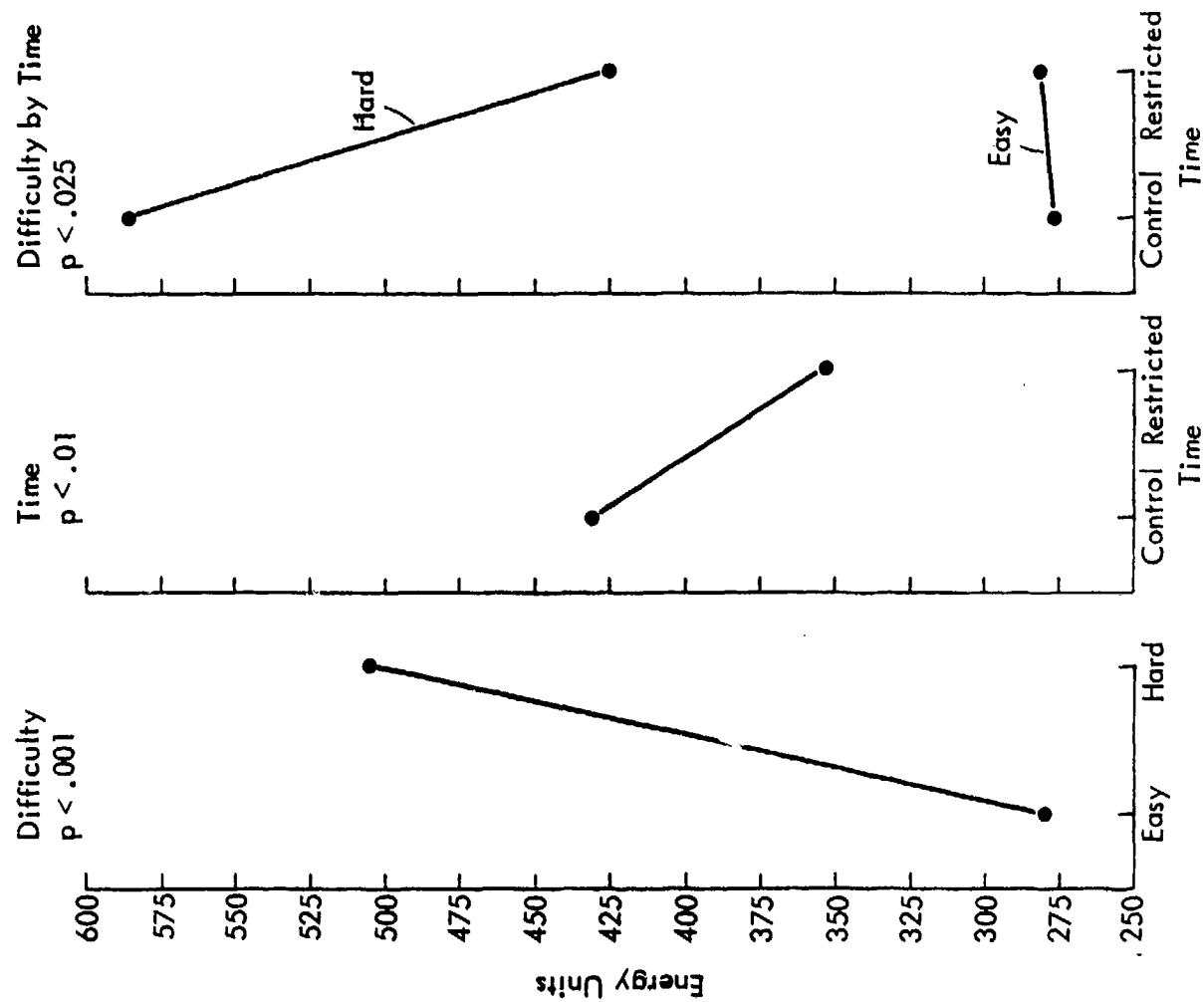


Figure 12 - Mean Shield Level on Entering Enemy Quadrants

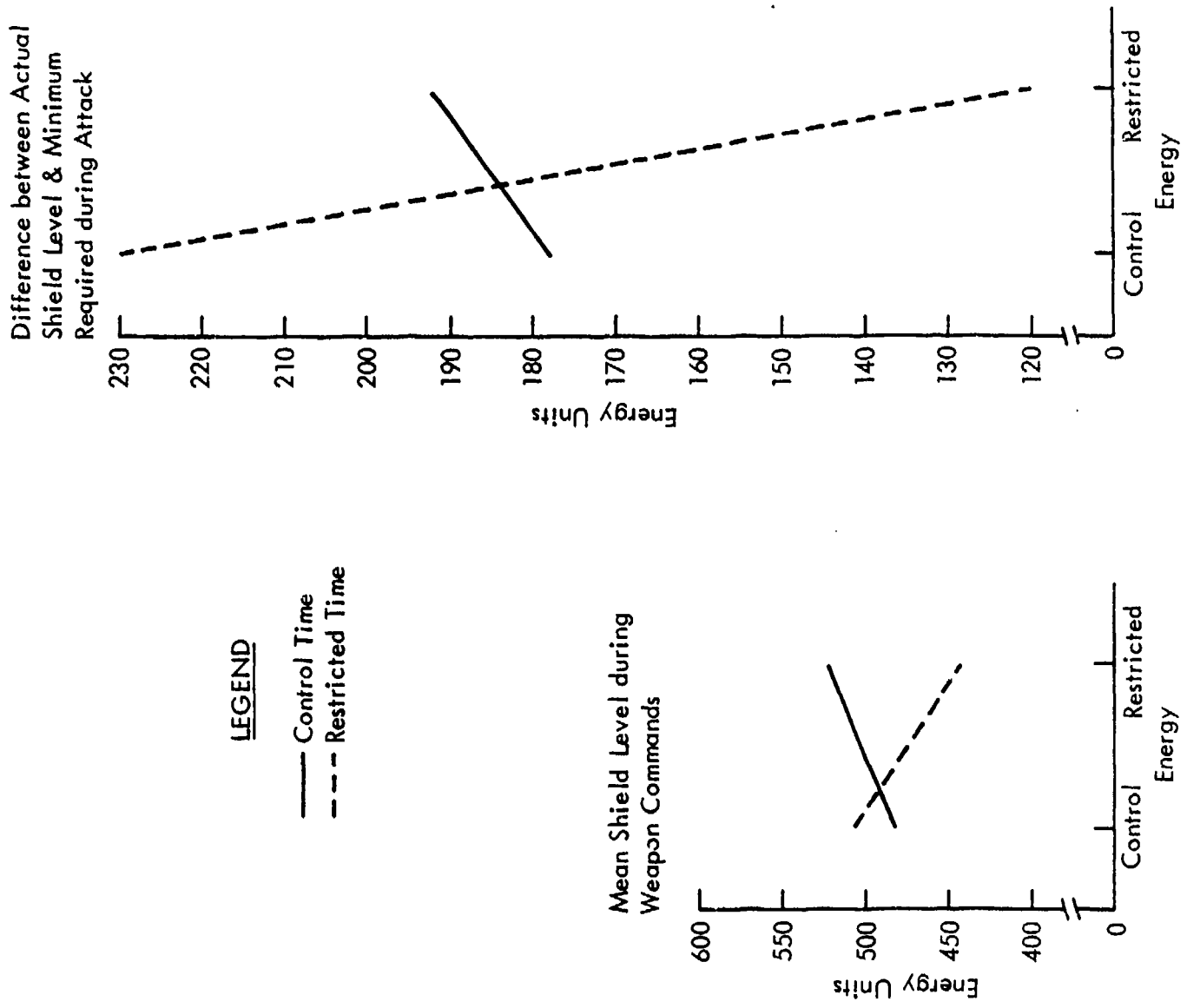


Figure 13 - Effects of Stressors on Risk-Taking Behavior During Attack Situations

As described above, a major impact of the stressors was in the area of risk-taking behavior. The stressors also had a significant impact on information processing measures. These effects are shown graphically in the following figures. Figure 14 presents the effects of the various experimental variables on the redundancy of information seeking behavior. Redundancy increases with difficulty level; however, under both time restriction and energy restriction, redundancy decreases. This effect is greater for hard than for easy galaxies. The combination of time and energy restrictions produced the least redundancy, and planned comparison revealed that the effect was significant only for the hard galaxies (control versus combined stressors,  $t = 4.59$ ,  $p < .001$ ).

Figure 15 shows the percent quadrants of new information per LRS. This measure differs from that described above in that it focuses on specified operational information-seeking, rather than total information seeking. More new information is gained per scan in easy than hard galaxies. Under time restrictions, subjects become more efficient in the sense that they gathered more new information per scan, and this effect was greater for hard galaxies. Planned comparisons between control galaxies and combined stressor galaxies revealed that more new information was obtained under combined stress conditions, but this effect was significant only for the hard galaxies ( $t = 2.89$ ,  $p < .02$ ).

Figure 16 presents the effects of the experimental variables on the percent noncontiguity of quadrant mapping. This variable measures the efficiency of the search strategy used by subjects to seek out the enemy. As noncontiguity increases, search efficiency decreases. Noncontiguity was greater in hard galaxies, and under energy restrictions. Energy restrictions, however, had no effect in easy galaxies. No main effect for time restriction was found. The 3-way interaction shown in Figure 16 is elucidated by planned comparisons between control and combined stress galaxies. Noncontiguity increased in combined stress galaxies, but this was limited to the hard galaxies ( $t = 4.15$ ,  $p < .01$ ).

Figure 17 presents the effects of stressors on the time subjects spend extracting information from their long range scans. No significant main effects were found. This is explained by the interactions shown in Figure 14. The effect of increasing difficulty alone is to reduce mean extraction time. The addition of time restriction reverses this effect; subjects spend more time extracting information in hard galaxies than in easy galaxies. When both time and energy are restricted, subjects spend more time extracting information during easy missions, but less time during hard missions. Replotting these data in the last panel of Figure 17 makes these interaction effects easier to see. The addition of stress conditions in easy level missions decreases the amount of time subjects spend extracting scan information. In hard level missions, stressors act to increase scan interaction time. The time stressor obviously has a marked effect on this measure. Under combined stress conditions, the differences in extraction time between easy and hard level missions is less compared to the imposition of either stressor alone.

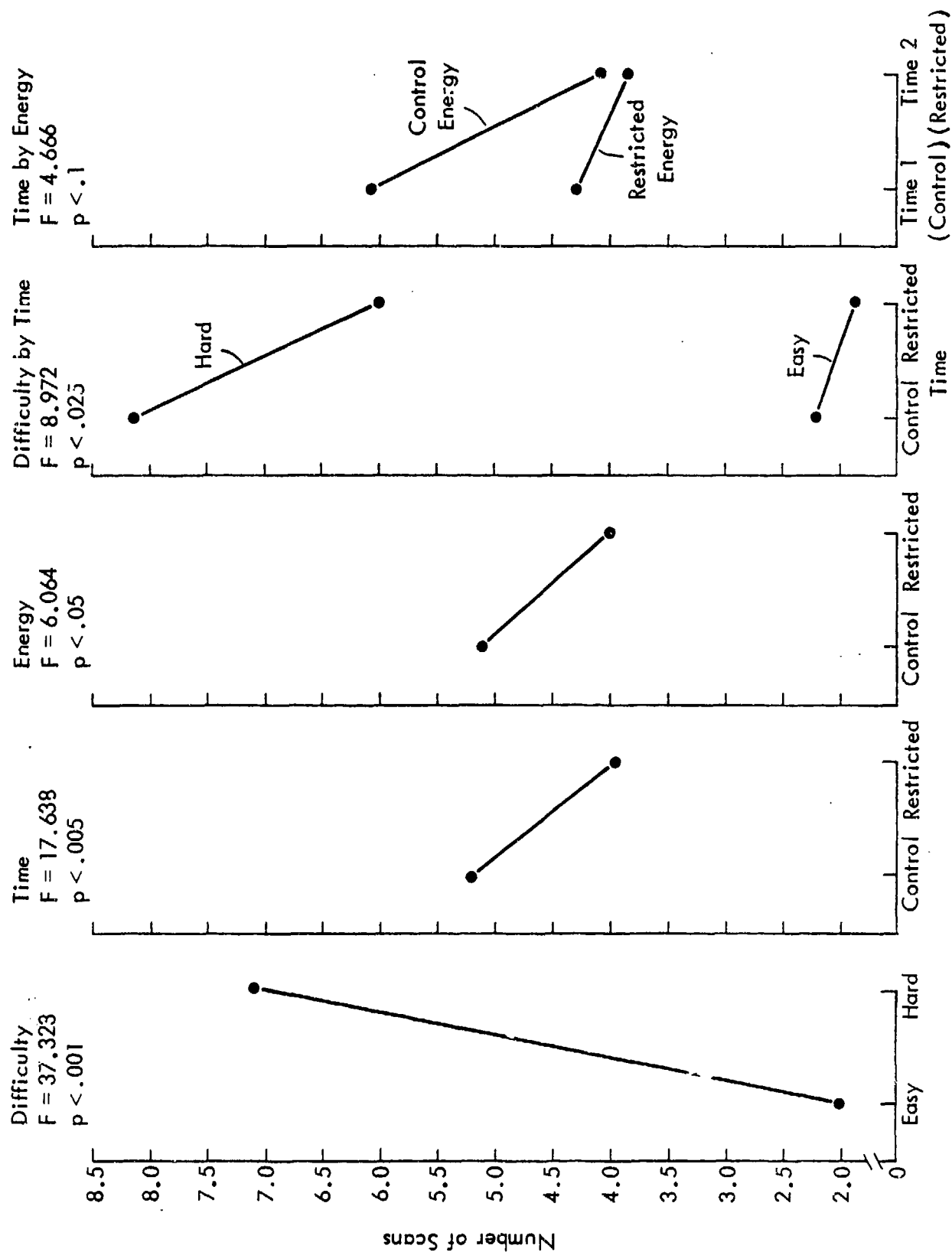


Figure 14 - Number of LRS and LRSII Commands With No New Information

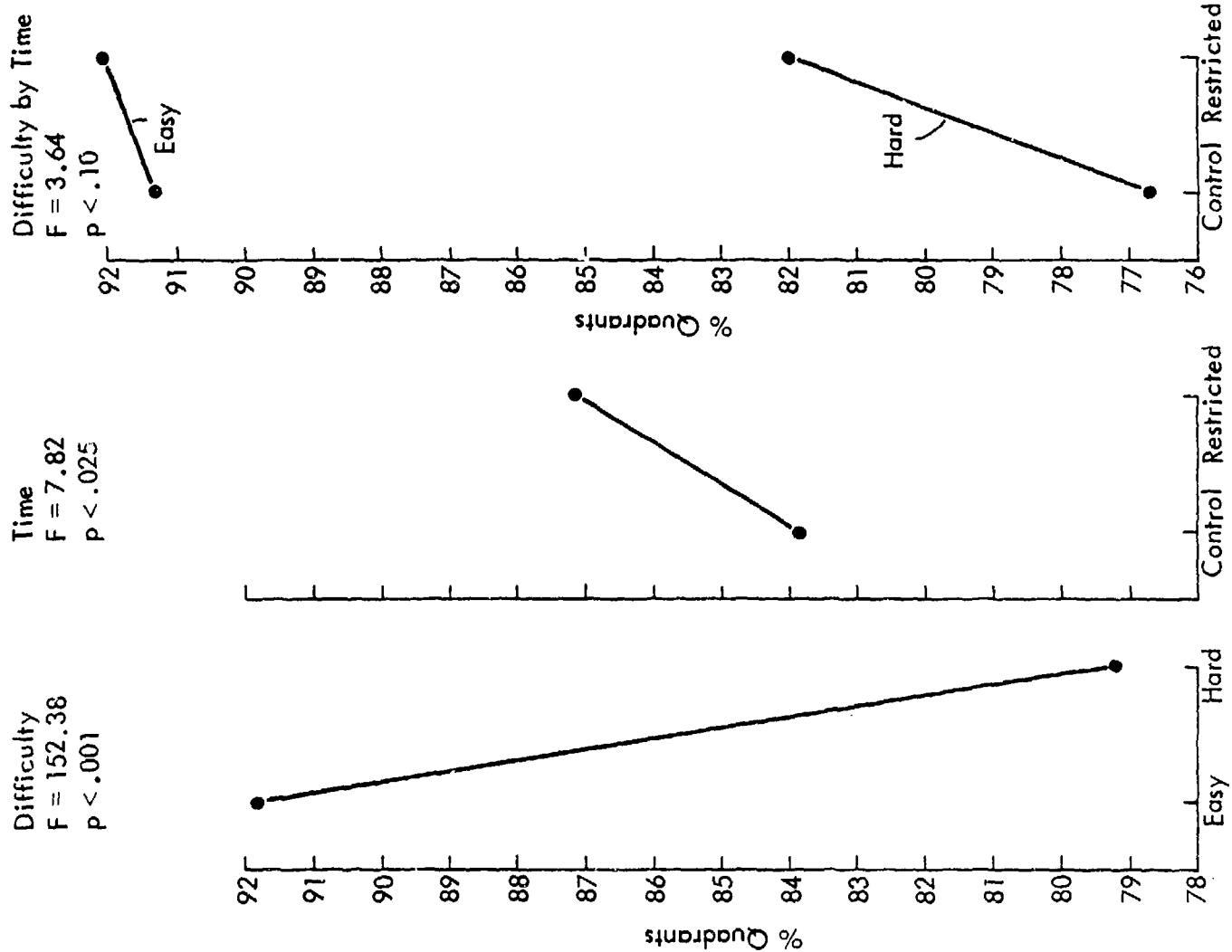


Figure 15 - Percent Quadrants of New Information per IRS

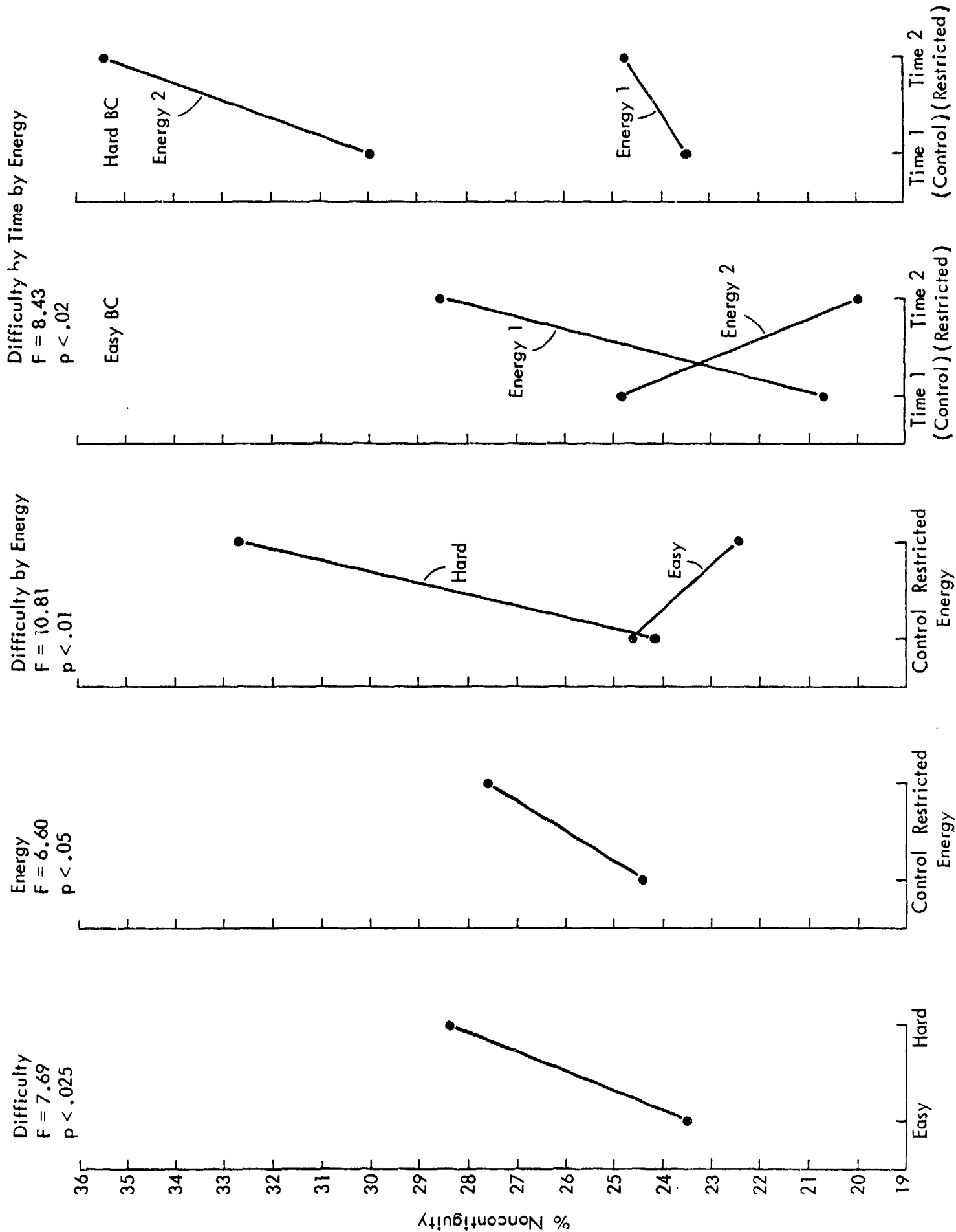


Figure 16 - Percent Noncontiguity of Quadrant Mapping

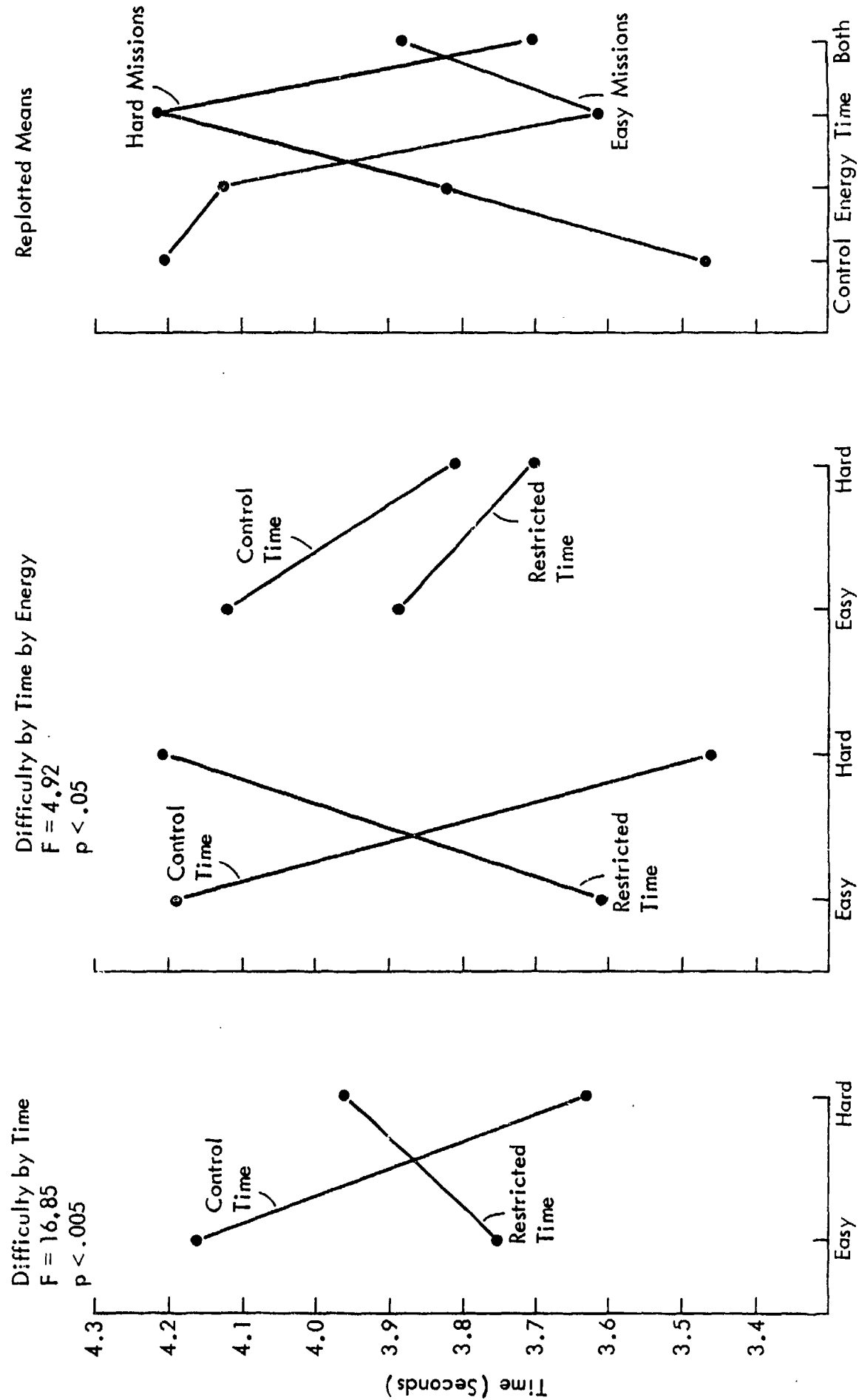


Figure 17 - Mean Time to Extract Information per IRS

5. **Discussion:** These findings suggest that the manipulation of time and energy restriction was effective in increasing ratings of stress and workload, and had marked consequences in terms of the numbers of Xenoids destroyed and the number of Venture cruisers lost. In the difficulty experiment, these highly practiced subjects destroyed all the enemy cruisers or all their missions. Under stress conditions, they failed to destroy 32 enemy cruisers, and lost their own cruisers on approximately one-third of all the missions conducted. These effects, as predicted, were most pronounced when the combined stress condition was in effect.

It should be noted that on most missions subjects did perform effectively under stress. However, when they did make an error, it tended to have large consequences. A primary response to the stress conditions was an increase in risk-taking behavior. Subjects decreased their shield protection on entering enemy quadrants. They issued significantly more weapon commands during attack situations without having adequate shield protection in the event that the attack failed. Finally, the mean difference between their actual shield levels and the absolute minimum required for adequate protection decreased significantly. These changes in performance meant that, in essence, subjects markedly decreased their margin for error. A small miscalculation could result in destruction of the Venture and/or failure of the mission.

These effects on overall performance measures and risk-taking behavior were most apparent when subjects performed missions under stress in easy galaxies. This was contrary to our original expectations. A possible explanation lies in the fact that, to an experienced subject, the difference between easy and hard galaxies is readily apparent at the beginning of the mission (triple versus single star density is easily recognized). Thus, subjects may have been motivated to take more care and expend more effort on hard missions.

The effects of the stressors were also apparent when information seeking and processing variables were examined. Under combined stress conditions in the hard galaxies, subjects increased their efficiency in the sense that they were able to obtain a greater percentage of new information per scan. This increase in the efficiency of information gathering, however, was offset by a decrease in the effectiveness of the strategy subjects used to search for the enemy. In order to gain new information, they developed a strategy of by-passing possible enemy locations, and leaving "holes" which they had to return to later in their search pattern.

The findings of this study are highly promising in regard to the use of STAR in future research in the area of stress and performance. The effects noted here were obtained with highly practiced subjects under basically minimal stress conditions of short duration. We would expect that under continuous performance demands or other experimental conditions resulting in increased stress, fatigue or workload, a number of STAR measures would prove to be sensitive indicators of decrements in cognitive function.

#### D. Evaluation of Test-Retest Reliability

At the conclusion of the training period, each subject performed two missions on each of two successive days to allow the evaluation of test-retest reliability of the measures. Means, standard deviations, skewness, and kurtosis, as well as Pearson's  $r$  between day 1 and day 2 means for each subject, were calculated for selected variables. Table 20 presents the results of this preliminary analysis. Evaluation of results suggested that either measurement procedure or calculations of variables were not optimal for all variables. For example, percent time on target in the standardized tracking task was neither normally distributed nor reliable because subjects had achieved extremely high levels of performance. Path size was therefore decreased for the difficulty and stress experiments, resulting in excellent normality of distribution. Similarly, reduction in time allowed to destroy Phantoms improved the distribution of reaction time measures.

During evaluation of these results, pilot tests for the difficulty experiment were being carried out. Xenoid dispersion was identified as a potential factor in determining mission difficulty. Since the galaxies used to evaluate test-retest reliability varied in Xenoid dispersion, we were concerned that the correlation coefficients obtained might underestimate reliability. Evaluation of reliability was therefore included in the design of the difficulty and stress experiments. Procedures and calculation algorithms were altered to improve normality of distribution, and a more extensive variable list was examined.

Subjects performed one "easy" and one "hard" mission in each of the two experiments. The galaxies for these missions were identical across the two experiments. Pearson's  $r$  between experiments, which were separated by 9 to 25 days, was calculated separately for easy and hard missions. Results are presented in Table 21. As would be expected from the long time between test and retest, the small number of subjects and the single-item nature of the reliability analyses, significant correlation coefficients were obtained for only about a third of the measures. Of particular interest are those nine variables for which correlations were significant in both hard and easy missions: number of LRSH with no new information, percent of commands used for LRSH, number of commands for LRSH, mean time between LRSHs, mean time to extract information from LRS, number of commands for search purposes, percent of commands used to attack, total number of commands, and percent commands for information seeking. Seven of the nine variables are associated with information processing and information seeking functions.

All four measures of perceptual accuracy and speed associated with torpedo use were reliable for either easy or hard galaxies, with greater reliability in general occurring in the easy galaxies. Memory measures based on errors of memory were more reliable than those assessing either percent correct recall or response time.

TABLE 20

PRELIMINARY EXAMINATION OF MEASURES FOR NORMALITY  
AND TEST-RETEST RELIABILITY

<u>Variable</u>	<u>Mean</u>	<u>S.D.</u>	<u>Normality<sup>a</sup></u>	<u>Pearson's <math>r</math></u>
<u>Overall Performance Efficiency</u>				
Mean time per enemy cruiser destroyed (sec)	34.6	6.4	Yes	0.68 <sup>b</sup>
Mean energy per enemy cruiser destroyed (sec)	200.1	32.3	Yes	0.11
<u>Psychomotor Skills</u>				
% time on target, tracking	97.3	2.1	S, K	0.50
Mean reaction time, Phantoms (sec)	0.53	0.10	K	0.85 <sup>b</sup>
% Phantoms destroyed	95.8	9.2	S, K	0.01
<u>Perceptual Accuracy</u>				
% quadrant entries as	97.0	4.0	S	0.42
Mean absolute torpedo course error	0.05	0.26	S, K	
<u>Memory</u>				
Overall % correct recall	74.6	9.2	Yes	0.50
<u>Information Processing</u>				
Mean quadrants information per LRS	148.0	22.7	K	0.45
Number redundant quadrants in LRS	86.4	22.4	K	0.45
Units of fuel for search	807.9	180.0	Yes	-0.37 <sup>b</sup>
Mean time to destroy enemy (from quadrant entry, sec)	8.1	4.0	Yes	0.62 <sup>b</sup>
Mean time to destroy supers (sec)	6.2	5.2	S, K	0.61 <sup>b</sup>
<u>Decision Making</u>				
Overall time per command (sec)	6.9	1.4	Yes	0.93 <sup>b</sup>
% correct weapon choice	86.0	9.1	K	-0.15

TABLE 20 (concluded)

<u>Variable</u>	<u>Mean</u>	<u>S.D.</u>	<u>Normality<sup>a</sup></u>	<u>Pearson's <math>r</math></u>
<u>Risk Taking</u>				
% phases with subminimum protection	22.9	14.3	S, K	0.36
<u>Subjective (1-9 scale)</u>				
Post stress	2.6	1.3	Yes	0.84 <sup>b</sup>
Post confidence	7.5	1.6	S	0.98 <sup>b</sup>
Post efficiency	7.4	1.5	S	0.96 <sup>b</sup>
Post workload	3.6	2.1	Yes	0.96 <sup>b</sup>
Post fatigue	2.8	1.6	K	0.91 <sup>b</sup>

a S = skewness > 1.0; K = kurtosis > 1.0.

b p < 0.05.

TABLE 21

TEST-RETEST RELIABILITY FOR STAR MEASURES

<u>Variable</u>	<u>Pearson's r</u>	
	<u>Easy</u>	<u>Hard</u>
<u>Overall Performance Efficiency</u>		
Mean time per enemy cruiser destroyed	0.47	0.34
Mean energy per enemy cruiser destroyed	0.23	0.06
<u>Psychomotor Skills</u>		
Tracking: Percent time on target	0.62 <sup>a</sup>	0.35
RMS error	0.43	0.28
Mean absolute error	0.56	0.25
Reaction Time: Mean reaction time	-0.52	0.30
% target detection	0.26	0.41
<u>Perceptual Accuracy and Speed</u>		
Torpedos: Number fired	0.84 <sup>a</sup>	0.07
% on target	0.81 <sup>a</sup>	0.13
Mean absolute course error	0.10	0.88 <sup>a</sup>
Navigation: % entries as intended	0.50	0.37
Mean execution time	0.46	0.50
<u>Memory Function</u>		
Overall: % correct recall	0.62 <sup>a</sup>	-0.01
Mean response time	0.40	0.55
Recent: % correct recall	0.31	-0.03
Mean response time	0.32	0.52
Mean post-pre response time	0.16	-0.09
Long-Term: % correct recall	0.19	-0.42
Mean response time	0.25	0.11
Errors of Memory: LRS, no new information	0.45	0.69 <sup>a</sup>
LRS, no new information	0.81 <sup>a</sup>	0.69 <sup>a</sup>
% commands, inadequate resources	0.68 <sup>a</sup>	0.50

TABLE 21 (continued)

Variable	Pearson's $r$	
	Easy	Hard
<u>Information Processing</u>		
Strategic:		
% commands LRS	0.23	0.82 <sup>a</sup>
% commands LRSH	0.90 <sup>a</sup>	0.80 <sup>a</sup>
Number of commands LRS	0.56	0.89 <sup>a</sup>
Number of commands LRSH	0.94 <sup>a</sup>	0.84 <sup>a</sup>
Mean time between LRS	0.59 <sup>a</sup>	0.72 <sup>a</sup>
Mean time between LRSH	-0.7	0.60 <sup>a</sup>
% noncontiguity of quadrant napping	0.10	0.01
Mean quadrants new information per LRS	0.53	0.50
Number of commands, search	0.78 <sup>a</sup>	0.85 <sup>a</sup>
Units of fuel, search	0.56	0.55
Tactical:		
Phaser calculation error	0.34	-0.03
% successful phaser commands	0.22	0.01
Mean time to calculate phaser, 1 enemy	0.26	0.18
Mean time to calculate phaser, 2 enemies	0.16	0.24
Mean time to calculate phaser, 3 enemies	0.60 <sup>a</sup>	0.13
Mean time to calculate phaser with supers	-0.24	-0.14
Mean time to calculate phaser without supers	0.82 <sup>a</sup>	0.22
Mean enemy destruction time	0.30	-0.13
% commands for navigate to attack	0.78 <sup>a</sup>	0.61 <sup>a</sup>
Fuel used for attack	0.58 <sup>a</sup>	0.34
Number of commands for navigate attack	0.22	0.30
<u>Decision Making</u>		
Number of Commands: Total	0.85 <sup>a</sup>	0.67 <sup>a</sup>
Number of Commands: Navigation	0.51	-0.11
% Commands for: Information seeking	0.86 <sup>a</sup>	0.68 <sup>a</sup>
Weapons	0.68 <sup>a</sup>	0.05
Defensive-resupply	0.25	0.62 <sup>a</sup>
Offensive-resupply	0.46	0.55
Damage control	NC	NC
Overall time per command	0.57	0.72 <sup>a</sup>
Navigations to different battle position	-0.16	0.08
Navigations to better battle position	-0.28	NC
% correct weapon choice	0.27	-0.52

TABLE 21 (concluded)

<u>Variable</u>	<u>Pearson's <math>r</math></u>	
	<u>Easy</u>	<u>Hard</u>
<u>Risk-Taking Behavior</u>		
Mean shield level when entering enemy quadrant	0.38	0.80 <sup>a</sup>
Mean shield level during weapon commands	0.72 <sup>a</sup>	0.44
% weapon commands, inadequate shields	0.57	0.24
Difference between actual and minimal required	0.45	0.02
Mean life support energy before resupply	0.55	0.68 <sup>a</sup>
Mean energy allocated to life support	0.45	0.68 <sup>a</sup>
Level of reserve energy and docking command	-0.19	0.30
Level of shield energy and docking command	NC	NC
<u>Subjective Ratings</u>		
Stress (post)	0.42	0.66 <sup>a</sup>
Confidence (post)	0.48	0.80 <sup>a</sup>
Efficiency (post)	0.34	0.77 <sup>a</sup>
Workload (post)	0.92 <sup>a</sup>	0.40
Fatigue (post)	0.27	-0.04

<sup>a</sup>  $p < 0.05$ .

NC = insufficient variance to calculate  $r$ .

The number of commands issued and the overall time per command were the most reliable of the decision-making variables. The lack of reliability in "percent correct weapon choice" may well be due to the fact that the algorithm used to calculate the measure did not take into consideration the available energy and armament resources. These items should be taken into account in the next version of STAR. Of the seven risk-taking variables, those involving maintenance of shield levels and maintenance of life support energy were most reliable. Although subjective ratings were highly reliable in the original study ( $r \geq 0.84$  for all items), no item was reliable across experiments for both easy and hard galaxies. It is not clear whether this is due to the nature of the instructions for the stress experiments, or to the relatively long time between test and retest.

These results are particularly encouraging when the item-by-item nature of the reliability tests is taken into account. Reliability of psychometric instruments and procedures increases as a function of the number of items included in the test. It would be expected, then, that a normalized score based on all items which address a specific concept (e.g., psychomotor skills) would be more reliable than any single item. Before determining reliability by variable class, empirical evaluation of the adequacy of the theoretical variable classes shown in Table 21 needs to be conducted. The existing data set contains too few subjects to perform the required principal components analysis. Subsequent work with STAR by this laboratory and others should provide the data necessary to more adequately evaluate test-retest reliability.

## V. CONCLUSIONS AND RECOMMENDATIONS

A primary goal of this project was to create a complex, cognitive performance task patterned after existing computer games, and incorporating unobtrusive, multiple measures of information processing, decision making and risk taking behavior. This goal was accomplished. In its current form, STAR will measure multiple aspects of an individual's performance, and it will do so without interfering with the ongoing performance activity.

The unobtrusiveness of the measurement procedures is a major, unique aspect of the task. This was accomplished by incorporating all STAR measures into the context of the task scenario. The standardized reaction time task is a good example of such incorporation. In a standard laboratory situation, a tone or a light is activated for a set number of trials, the subject responds as quickly as possible, and the experimenter records performance values. The subject knows that this is testing his or her ability to react quickly and also that someone is evaluating her or her performance during the test. The test is also an isolated unit of performance, not connected to, or integrated with, any larger task of direct interest to the subject.

STAR also contains a standardized reaction time task. It is exactly the same type of task used thousands of times previously in research laboratories. The basic difference between STAR and previous tasks

lies in the context in which the reaction time task is presented. In STAR, the subject will encounter five phantom Xenoid cruisers on any mission. A Phantom appears immediately after entering an enemy quadrant, and the subject has 1 sec to protect the VENTURE from enemy fire by depressing a key on the terminal which activates the defense shields. Thus, on any mission in STAR, there are standard reaction time trials which are presented at specific points in the mission and there is a standard response. What is different is that the reaction time task is incorporated into the larger context of the mission, no one is obviously evaluating this particular aspect of performance, and the consequences of performance error are now personally relevant to the subject.

Almost all STAR measures are incorporated into the task context in a similar fashion, and this feature of the measures allows analysis of integrative rather than isolated, episodic performance. For example, it is much more relevant to know that task difficulty or stress affects the ability to detect and respond to meaningful target stimuli, than it is to know that reaction time changed so many milliseconds. Thus, one of the major advantages of STAR is its capability to measure and evaluate integrated components of real-time, ongoing behavior without, at the same time, influencing or interfering with the behavior being measured.

The scenario selected for STAR is that of a futuristic war, and the task incorporates elements of current day computer games. This was a deliberate choice for the prototype, since these elements tend to enhance task involvement and maintain high levels of motivation. Task elements, however, could be created and combined in a number of different ways to form other contexts and scenarios. For example, the task could just as easily have been set in the context of a ground forward area observer flying above a battlefield. Similarly, the measures incorporated into STAR presently encompass a number of cognitive performance areas. This was a deliberate choice also, since we envisioned the task as a general purpose research tool. However, it is also quite possible to "tailor" the measures and the context such that they are directly relevant to specific real world functions and duties. Thus, one of the primary points demonstrated by the development of STAR, is that elements and measures can be created and combined to form a contextual research tool with the capability to unobtrusively evaluate components of complex performance.

There are some major differences between a task such as STAR and the typical simulation exercises currently available. Simulation exercises generally evaluate the performance of a group, usually at the battalion level, conducting a multi-day, computerized land battle over known terrain. The group is not learning to perform a task; rather they are practicing what they have already been highly trained to do. Information about the consequences of the actions directed against the enemy is often not known until some time after the action is taken (e.g., holding a bridge). In contrast, STAR assesses individual performance. The mission duration is considerably shorter than a simulated battle. The terrain is not known. Individuals have to learn to perform STAR, and the consequences of actions taken are known almost immediately.

These contrasts serve to highlight the different purposes STAR has been designed to accomplish, compared to those for which simulations are designed. For example, one major purpose of simulations is to provide comparisons across different battalion command groups in the performance of a fairly standardized battle scenario. STAR, on the other hand, is designed to address questions of a more basic nature concerned with the impact of various stress and workload conditions on relevant components of cognitive efficiency. Examples of such questions are as follows.

- \* If highly trained individuals are required to perform a well-learned, highly motivating, complex task continuously for long periods of time, what happens to their normal levels of cognitive efficiency? If difficulty level and crisis conditions are systematically introduced and varied, can individuals modulate their level of effort so as to maintain efficiency in the face of changing demands?
- \* If efficiency decreases under the above conditions, to what is the decrease due? In other words, what are the types of errors made, and what were the decision paths that led to such errors?
- \* If an individual is required to perform a complex cognitive activity while wearing chemical protective gear, how does this change in and of itself influence performance?
- \* Similarly, what is the effect of the injection of protective drugs such as atropine, or the injection of combinatorial drugs such as pyridostigmine and oximes, on cognitive function?
- \* Most research has focused on college students; however, the majority of influential decision makers are over 35 years of age. What is the impact of age and experience on the ability to perform complex cognitive activities under stress?
- \* How do people go about learning to perform a complex cognitive task such as STAR? In other words, how is new knowledge incorporated into older existing knowledge structures? How can instructional and training strategies be improved to speed up the process and make it more efficient? How can training information be better "packaged" to eliminate information overload?
- \* How do people actually make complex cognitive decisions in a well-learned activity? What is the information they use, the sources they tap, the organizational structures they impose, and how do they weigh the validity and importance of incoming information under conditions of ambiguity, overload or stress?
- \* Why do some people under stress conditions seem able to deal effectively with multiple channels of information and interspersed, complex decision making, and others do best by performing in a "one-thing-at-a-time" mode? Can a person's

"style," or mode of cognitive operation be predicted beforehand?

- \* Can physiological and biochemical measures which would lead to increased understanding of human information processing and complex decision making be incorporated into STAR. For example, can the brain's event related potential be used to aid in determining whether an error was due to failure to perceive a target stimulus, failure to recognize its importance, or failure to respond appropriately?

The above questions, as well as others, all constitute incapsulated descriptions of viable areas of future research. The answers to such questions would have practical relevance in the areas of chemical defense, selection and training, and in the measurement and evaluation of C<sup>3</sup> operations. It should be noted, however, that STAR is a prototype, a beginning, of the kind of research tools we feel will be coming into more frequent use in the future. The experiments presented here were conducted with the idea of providing preliminary data on major aspects of the task, and on some of the projected uses of the task. While they are promising, it is important to realize that these data were collected on a total of only 22 subjects. More subjects need to be evaluated, and in particular, more statistical analyses of the measures need to be conducted. Analysis of larger data sets would provide the means to empirically evaluate the adequacy of the measurement variables and constructs currently available in STAR. We strongly recommend that this be given priority in future research efforts. Appendix B presents a more detailed description of some suggestions for future research.

APPENDIX A

STAR TRAINING PROTOCOL

Appendix A  
to  
Annual and Final Report

Task Validation for Studies on Fragmented Sleep  
and Cognitive Efficiency Under Stress  
Contract No. DAMD17-80-C-0075

STAR TRAINING PROTOCOL

The training protocol: The basic protocol consists of a sequence of three types of training sessions.

\* Initial group orientation session: This is the first session of the protocol. Its purpose is to introduce the task to multiple trainees. The purpose of the research activity is explained, informed consent is obtained and training schedules are clarified with each trainee. Each trainee is provided with a copy of the training manual, a pad of paper and several pencils. The group is told that in this session they will be asked to read the manual, see a short demonstration of STAR, work the exercise problems, and if time allows, practice the tracking task incorporated into STAR. This session requires approximately 3 hr to complete, including a 5 to 10 min mid-session break.

When introducing the training manual, the instructor should emphasize the following:

- \* The trainee will be able to keep the manual for reference purposes during the training program.
- \* Trainees will have 45 min to read the manual.
- \* The purpose is simply to gain familiarity with the task and the procedures involved and not to learn each operation in detail.
- \* The trainees will see a demonstration of STAR following their reading, and this demonstration will help clarify the material presented in the manual.
- \* As they read the manual, they should ask questions about anything that is unclear to them.
- \* They should not start on the training exercise problems.

After reading the manual, trainees are shown a brief 15 min demonstration of STAR. The demonstration should start by showing the trainees how to operate the computer terminal. The demonstration mission is then started, and the trainee is shown how to:

- \* Interpret the initial mission briefing information presented.

- \* Pilot the shuttlecraft out to the HUB.
- \* Answer the questions asked by the HUB commander.
- \* Operate the command and control system onboard VENTURE.
- \* Recognize the different types of scan information presented.
- \* Use both types of weapon systems available.
- \* Defend against a Phantom Xenoid attack.
- \* Navigate and search for the enemy.
- \* Receive a Mission Debriefing Report after mission completion.

The trainees then receive a 5 to 10 min break. The remainder of the session is devoted to working the exercise problems. The exercise problems present typical situations the trainee will encounter during actual task performance. They focus on four basic skills: navigation, photon torpedo fire control, phaser fire control, and tactical decision making. The 40 problems, 10 in each of the above skill areas, are arranged in increasing order of complexity within each skill area. Answers to all problems are presented at the end of the exercise problem section of the manual. Trainees should work the first problem, check the accuracy of their answer, work the second problem, and so on. Questions should be encouraged, and the instructors should be highly familiar with the problems.

If a trainee finishes the problems and still has at least 20 min left in the session, he or she can go on to practice the tracking task. The special purpose program used to train tracking skill is called PATH. The source listing for PATH, an operational copy of the program, and instructions are included in the attachments to this report.

The PATH program is designed as an automated tracking-teaching program. Once the instructor calls up the program and sets in the performance progression parameters, skill acquisition is completely under the control of the trainee. Training starts at the easy level of difficulty. As trainees meet pre-set performance criteria, bells go off, the trainee is congratulated by the computer, and the next higher level of difficulty is automatically activated. The trainee then works at that level until criterion is reached, and so on. The program stops after criterion performance is reached at the most difficult level. The vertical movement of the tracking path is controlled by the integration of 3 sine waves. The factor that controls task difficulty is the width of the path the trainee must keep the cursor inside. The easy difficulty level has a path width of 14 characters on an 80-character wide display screen. The highest difficulty level has a path width of 8 characters. This is the level actually used in the STAR task. PATH training ends when a trainee can stay in the path at least 80% of the time on each of three successive 30-sec. trials.

It should be noted that the instructor does not have to continuously monitor trainee performance on PATH. Once the parameters are set in and the task is demonstrated to the trainee, the instructor can attend to other things. The program in interaction with the trainee, will repeatedly present the task, provide performance feedback to the trainee after each trial, change criteria levels, and collect data on each trial (% time in path, RMS error, X absolute error). The bell that is activated as each criterion level is achieved is audible over open intercom circuits and the instructor can monitor performance acquisition from anywhere in the vicinity. This type of intensive tracking practice should not be provided for longer than a 20 min period. Longer periods tend to result in the build-up of forearm and wrist muscle fatigue, and this interferes with the acquisition of tracking skill.

\* Individual training with performance coaching: At the start of subsequent sessions of the training protocol, trainees first practice the tracking task until either 20 min elapse or they reach criterion at the highest level of difficulty. Following tracking practice, the instructor coaches the trainee through the first practice mission. This mission is directly identified as a practice mission to the trainee. He or she is told that the performance values obtained will not "count". This procedure helps to reduce the performance anxiety that is often present. The trainee is seated before the computer keyboard/ display and the instructor is seated to the side. The trainee operates the computer and the instructor's role is to provide help and information, but not to take over and lead the trainee through the mission. It is good practice to let the trainee make several mistakes, even to the point of destroying the VENTURE, and then to go over why it happened and how it could have been avoided. The instructor should try to create an informal, friendly, and helpful atmosphere, while at the same time, indicating clearly that it is the trainee who is expected to learn to make the necessary decisions.

Prior to the start of the mission, the instructor should again familiarize the trainee with the operation of the computer keyboard. The instructor should also point out that copies of the navigation and attack vector system (Table 2, page 13 of the training manual) and the fraction to decimal conversion table (page 26 of the manual) are attached to the display terminal for use by the trainee. The most important point to make clear to the trainee concerns the calculation of navigation and torpedo courses, and phaser payloads. In working the training problem during the first session, trainees were free to use paper and pencil to make their calculations. In this session, and in all subsequent sessions, all numerical calculations have to be made without the aid of paper and pencil. The reason for this is that a number of information processing and decision making measures derived from STAR are time dependent. The use of paper and pencil by some trainees and not others reduces the adequacy of these measures to accurately reflect these aspects of cognitive function.

The instructor should expect that many trainees will have difficulty remembering the details of particular operations. Specific advice is often required in the following areas:

- \* Where to start a search pattern when first entering the galaxy
- \* How to actually use the scan information
- \* How to navigate
- \* What levels of energy to allocate to various systems
- \* Calculation of phaser payloads in multiple Xenoid quadrants
- \* How to "spot" a better battle position

Depending on the trainee's grasp of the STAR task, coaching can continue for up to three missions. Coaching activities in missions two and three should become progressively less directive. Once the instructor feels reasonably assured that the trainee appears to understand the fundamentals of the task, the final training phase can be initiated. This phase is described below.

\* Training to performance criteria: At the start of the training to performance criteria (TPC) missions, the following information is provided to trainees. Starting with this mission, the trainee will be given the lowest command rank. His or her goal will be to meet progressively more difficult time and energy promotion criteria until the trainee can perform at the highest level of VENTURE Captain Level 1. The trainee will have a total of eight missions to achieve the first promotion criterion (destroy all Xenoids and not lose VENTURE in the process). If criterion is not reached, the trainee will be discontinued from further training.

The trainee will be alone in the performance room during the mission. The instructor will be monitoring the mission on the "slave" monitor in the control room. If the trainee needs advice, or if the instructor sees a problem developing, they will be able to communicate over the intercom. At the end of each mission, the trainee and instructor will review the Mission Debriefing Report together and discuss the mission.

The instructor sets in the appropriate mission and subject identification parameters, starts STAR, and leaves the performance room. The mission is monitored from the control room. The instructor should carefully monitor that the trainee is performing the command sequences correctly, and is attempting to answer the HUB Commander's questions accurately. Trainees often tend to not realize how much time they are taking in making and executing command decisions. After they have conducted a successful action against one enemy concentration, often they will relax and congratulate themselves, instead of immediately searching for additional enemy concentrations. In order to reach the highest performance level (VENTURE Captain Level 1) trainees have to be both fast and accurate. It is appropriate for the instructor to "come on" the intercom and tell trainees to speed it up, or to enter the commands correctly. These interactions should be brief and concise.

Similarly, the instructor needs to be aware that some trainees develop a dependence on the instructor to tell them what to do next. It is appropriate for the instructor to answer trainee questions regarding factual information (e.g., a torpedo hit costs how much?) or procedures (e.g., do I enter quadrant or sector coordinates?). It is not appropriate to continuously answer tactical or strategic questions (e.g., should I search quadrant 3,2 or 4,7?). A good instructional response to such questions is to point out that the Xenoids may be monitoring communications, and the instructor cannot answer that question. This gets the idea across to the trainee that certain questions are inappropriate, and that decision making is the responsibility of the trainee.

In the initial TPC missions, the instructor needs to use judgment concerning the extent to which he or she interacts with the trainee over the intercom. The basic goal of the training is to enable trainees to move quickly up to Level 1 performance. Individuals differ in their ability to do this, and in the speed with which they can progress through the various performance levels. Mission performance needs to be closely monitored, as does the response of the trainee to victory and defeat.

After each mission, the trainee receives a computer-generated Mission Debriefing Report and is informed if he or she will be promoted, demoted or will remain at the present rank. This report, together with the comments of the instructor, constitute an important element in the training program. It is during this time that the instructor can discuss in depth the areas where the trainee is doing well and where more work is needed. Training is concluded for an individual when he or she can perform at the criterion set for VENTURE Captain Level 1 over three sequential missions.

APPENDIX B

SOME SUGGESTED AREAS FOR FUTURE RESEARCH

Appendix B  
to  
Annual and Final Report

Task Validation for Studies on Fragmented Sleep  
and Cognitive Efficiency Under Stress  
Contract No. DAMD17-80-C-0075

SOME SUGGESTED AREAS FOR FUTURE RESEARCH

Under this contract, a computer gaming approach was used to create a new type of automated performance task (STAR). STAR unobtrusively measures multiple components of complex human performance, including psychomotor skills, information processing, memory, decision making, risk-taking, and subjective status. A training manual and protocol have been developed, training criteria established, and the effects of changes in task difficulty and crisis conditions on performance of the task assessed.

STAR can now serve as a test bed which will provide operationalized measures of parameters relevant to complex human performance. The efficiency with which STAR could be used as such a test bed, however, could be improved by mathematically combining specific measures to yield an overall score for specific classes of cognitive functions and performance parameters. At present, specific measures have been grouped together on theoretical grounds. Although a large amount of mission data is currently available for analysis, these missions were performed primarily by 14 different individuals. The data set is therefore too small to perform the types of analyses necessary to yield empirically verified scores for specific functions. Future studies should focus on generating a sufficiently large data set for the application of multivariate statistical procedures.

Several specific lines of research are suggested by our experience with the task.

1. Fatigue and recovery from fatigue: Recent advances in weapons research are changing established concepts of warfare. Continuous and sustained operations are now technically feasible. The impact of fatigue associated with such operations on relevant components of complex human performance needs to be further delineated. Of particular importance are the identification of likely areas of human error, the effects of individual differences on the ability to maintain adequate performance, and the pattern of recovery function.

2. Complex task learning: STAR is a complex task to learn. It involves interacting with the computer, extracting information from displays, learning multiple command sequences, anticipating and recognizing situations requiring specific actions, and making complex decisions often on the basis of incomplete or ambiguous information. These characteristics of STAR are highly similar to those encountered in a number of current high-technology, military specialties. With appropriate modification, STAR could provide an excellent tool for the investigation of theoretical issues concerned

with complex skill acquisition, and for testing new approaches to the training and/or retraining of personnel. For example, in the introduction to the stress study described in this report, we described the marked impact of changing one task variable (resource availability) on task learning and relearning. Future research focusing on learning and training issues could prove valuable in a number of practically relevant areas of the current military environment.

3. Formation and functioning of crews, teams, and units: As presently designed, STAR is performed by a single individual. Pilot observations in our laboratory, however, indicate that the functions performed by the individual could be divided naturally between individuals (S1-S4). This would allow a small group to perform the present STAR task with only minor modifications required. This type of group performance configuration could be used to investigate the development of small unit formation; group bonding, interaction and motivation; leadership characteristics and efficiency; and group function under stress.

4. Age and sex differences in complex task performance: In the present study, STAR was performed by college students of both sexes. Our findings indicated that sex did not make a difference in overall task performance or speed of acquisition. However, the processes and procedures men and women used to reach their performance goals did seem to differ. This suggests that STAR could be used to examine differences in cognitive processing and decision making between men and women, within a complex task which each can perform equally well.

Similarly, STAR is a task requiring skill at information processing and decision making; it is not a task that is dependent on fast reaction time or rapid visual motor coordination. Thus, it can be used in future research to evaluate the advantages or disadvantages of age and experience on complex task performance under a variety of conditions. It should be noted that relatively little relevant research has been directed at the age groups who in the real world actually make most of the influential decisions. In the military environment, it is these groups who will be processing information and making the critical decisions under sustained operation conditions and/or under various environmental stressors.

APPENDIX C

EVALUATION OF STAR AS AN EXPERIMENTAL TOOL

Appendix C  
to  
Annual and Final Report

Task Validation for Studies on Fragmented Sleep  
and Cognitive Efficiency Under Stress  
Contract No. DAMD17-80-C-0075

EVALUATION OF STAR AS AN EXPERIMENTAL TOOL

The goal of this project was to develop what we have come to think of as the prototype for the "next generation" of research tools to evaluate complex cognitive performance. This goal was accomplished. The prototype task (STAR) provides unobtrusive multiple measures of complex performance within the context of a highly-motivating game scenario. The purpose of this appendix to the final report is to provide expert opinion about how this new approach differs from older, more traditional approaches, and what our observations during the performance of this project have indicated to us that might be of value to others. It is opinion, but this opinion is based on a combined total of approximately 30 years of staff research experience in relevant areas.

STAR is without a doubt the most motivating task we have ever seen used in a research laboratory. During the recruitment of subjects for this project, word-of-mouth advertizing from subjects to their friends resulted in many more volunteers calling in than we could handle. During the training phase, we budgeted 18 hr as the initial total training time. If subjects did not achieve Level 1 in that period of time, they often wanted to come in on their own time and for no payment just so they could have the personal satisfaction of being a Level 1 captain.

The task is not only generally motivating, it is highly involving during the performance of missions. Subjects perform STAR alone in a small room in the laboratory; however, we maintain an open intercom connection from that room to another part of the laboratory. During missions, subjects shout, curse, and groan when they make a mistake; this has caused visitors to the lab to wonder how we are treating subjects! We have also had to caution some subjects not to pound the computer keyboard so drastically during missions.

The motivational properties of STAR are a distinct aid in performance assessment. First, all the volunteers in this project were fully informed of what we were doing and what we wanted to accomplish. In other words, they knew we were evaluating their performance. However, once they became involved in the mission, and we were no longer physically present in the room, this aspect of the project faded into the background. The evaluation process no longer interfered with the performance.

Second, a primary problem in past research has been how to maintain high task involvement and sustained performance over long periods of time. This does not seem to be a problem with STAR. The only indication of somewhat reduced interest we have seen, happened when highly proficient

Appendix C  
to  
Annual and Final Report

Task Validation for Studies on Fragmented Sleep  
and Cognitive Efficiency Under Stress  
Contract No. DAMD17-80-C-0075

EVALUATION OF STAR AS AN EXPERIMENTAL TOOL

The goal of this project was to develop what we have come to think of as the prototype for the "next generation" of research tools to evaluate complex cognitive performance. This goal was accomplished. The prototype task (STAR) provides unobtrusive multiple measures of complex performance within the context of a highly-motivating game scenario. The purpose of this appendix to the final report is to provide expert opinion about how this new approach differs from older, more traditional approaches, and what our observations during the performance of this project have indicated to us that might be of value to others. It is opinion, but this opinion is based on a combined total of approximately 30 years of staff research experience in relevant areas.

STAR is without a doubt the most motivating task we have ever seen used in a research laboratory. During the recruitment of subjects for this project, word-of-mouth advertizing from subjects to their friends resulted in many more volunteers calling in than we could handle. During the training phase, we budgeted 18 hr as the initial total training time. If subjects did not achieve Level 1 in that period of time, they often wanted to come in on their own time and for no payment just so they could have the personal satisfaction of being a Level 1 captain.

The task is not only generally motivating, it is highly involving during the performance of missions. Subjects perform STAR alone in a small room in the laboratory; however, we maintain an open intercom connection from that room to another part of the laboratory. During missions, subjects shout, curse, and groan when they make a mistake; this has caused visitors to the lab to wonder how we are treating subjects! We have also had to caution some subjects not to pound the computer keyboard so drastically during missions.

The motivational properties of STAR are a distinct aid in performance assessment. First, all the volunteers in this project were fully informed of what we were doing and what we wanted to accomplish. In other words, they knew we were evaluating their performance. However, once they became involved in the mission, and we were no longer physically present in the room, this aspect of the project faded into the background. The evaluation process no longer interfered with the performance.

Second, a primary problem in past research has been how to maintain high task involvement and sustained performance over long periods of time. This does not seem to be a problem with STAR. The only indication of somewhat reduced interest we have seen, happened when highly proficient

analysis of a larger data set) but the current measures do seem to provide a reasonable good "picture" of what happens during a mission. This could be markedly improved through further analyses of additional data.

In summary, what we have tried to do in this section is to provide an informal overview of our impressions of STAR. It is a unique task in many ways, and it should prove valuable in a number of areas of future research.

## DISTRIBUTION LIST

12 copies	Director Walter Reed Army Institute of Research Walter Reed Army Medical Center ATTN: SGRD-UWZ-C Washington, DC 20012
4 copies	Commander US Army Medical Research and Development Command ATTN: SGRD-RMS Fort Detrick, Frederick, MD 21701
12 copies	Defense Technical Information Center (DTIC) ATTN: DTIC-DDA Cameron Station Alexandria, VA 22314
1 copy	Dean School of Medicine Uniformed Services University of the Health Sciences 4301 Jones Bridge Road Bethesda, MD 20014
1 copy	Commandant Academy of Health Sciences, US Army ATTN: AHS-CDM Fort Sam Houston, TX 78234